



Pattison, Richard John (2008) *Characterisation of the PQ-loop repeat membrane protein family in arabidopsis thaliana*.
PhD thesis.

<http://theses.gla.ac.uk/346/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

**Characterisation of the PQ-Loop
Repeat Membrane Protein Family in
*Arabidopsis thaliana***

by

Richard John Pattison

Thesis submitted for the degree of doctor of philosophy

Division of Biochemistry and Molecular Biology,

Institute of Biomedical and Life Sciences,

University of Glasgow

April, 2008

Contents

Title	i
Contents	ii
List of figures and tables	vi
Terminology	viii
Abbreviations	ix
Acknowledgements	x
Abstract	xi
Chapter 1: Introduction	1
1.2 Cystinosin (CTNS)	1
1.2.1 Biological importance of cystine	3
1.2.2 CTNS-like function in plants	4
1.3 MPDU1 and ERS1	5
1.3.1 Man-P-Dol and Glc-P-Dol dependent glycosylation	5
<i>1.3.1.1 Synthesis, Reorientation and Utilisation of Man-P-Dol and Glc-P-Dol</i>	6
<i>1.3.1.2 Mutations in A. thaliana affecting Man-P-Dol synthesis</i>	7
1.3.2 N-linked glycosylation	8
<i>1.3.2.1 Lipid-linked oligosaccharide synthesis</i>	8
<i>1.3.2.2 Chemical inhibitors of LLO synthesis</i>	10
<i>1.3.2.3 N-Glycan oligosaccharyltransferase mutants</i>	11
<i>1.3.2.4 N-Glycosylation and the calnexin-calreticulin cycle</i>	12
<i>1.3.2.5 N-Glycan modification</i>	14
1.3.3 O-linked Glycosylation	14
1.3.4 C-linked Mannosylation	15
1.3.5 GPI anchoring	16
<i>1.3.5.1 GPI anchor biosynthesis</i>	16
<i>1.3.5.2 GPI anchor biosynthesis mutants</i>	17
1.3.6 Importance of Man-P-Dol and Glc-P-Dol for viability	18
1.3.7 Isolation of MPDU1/LEC35	20
<i>1.3.7.1 Isolation of lec35 mutants and analysis of N-linked glycosylation</i>	22
<i>1.3.7.2 Analysis of GPI anchoring and C-mannosylation in lec35 mutants</i>	23
<i>1.3.7.3 Possible functions of LEC35</i>	24
<i>1.3.7.4 Genetic interaction between LEC35 and GPT</i>	25
<i>1.3.7.5 LEC35 homologues in other species</i>	26
<i>1.3.7.6 Prediction of a lec35-like mutant phenotype</i>	26
1.3.8 ERS1	27
<i>1.3.8.1 ers1Δ mutant and complementation by CTNS</i>	28
1.4 STM1	29
1.4.1 G-protein signalling	29
1.4.2 Isolation of STM1	29
<i>1.4.2.1 Functional characterisation of STM1</i>	30
<i>1.4.2.2 STM1 as a GPCR?</i>	31
1.4.3 G-protein signalling in plants	32

1.4.3.1 GPCRs in plants	32
1.4.3.2 Predicted phenotype of an <i>stm1</i> -like mutant in plants	34
1.5 Objectives and thesis outline	35
 Chapter 2: Materials and Methods	 36
2.1 Materials	36
2.2 General Methods	36
2.2.1 Plant Material and Growth Conditions	36
2.2.2 RNA extraction and cDNA synthesis	38
2.2.3 Polymerase chain reaction (cloning, zygosity testing and expression analysis)	39
2.2.4 Molecular cloning and plasmid construction	43
2.2.5 <i>E. coli</i> transformation and plasmid minipreps	47
2.2.6 <i>A. tumefaciens</i> transformation	49
2.2.7 <i>A. thaliana</i> stable transformation	50
2.3 Methods for chapter 3: alignments and sequence analysis	51
2.3.1 DNA and amino acid sequences	51
2.3.2 Alignment and phylogenetic tree construction	52
2.3.3 Hydropathy plots	52
2.3.4 Blast and Pfam searches	52
2.3.5 Gene duplication testing	53
2.3.6 Analysis of promoter regions	53
2.4 Methods for chapter 4: expression profile and sub-cellular localisation	53
2.4.1 Plant growth conditions, viral and wounding treatments	53
2.4.2 Promoter GUS fusions – transformant selection and histochemical staining	55
2.4.3 <i>AtPQL</i> -GFP fusion construct expression and confocal microscopy	56
2.4.4 Public microarray expression database queries	57
2.5 Methods for chapter 5: functional characterisation of the <i>AtPQL</i> family	58
2.5.1 Yeast strains and media	58
2.5.2 Yeast transformation and growth tests	60
2.5.3 T-DNA insertion mutant isolation	61
2.5.4 Over-expresser mutant isolation	62
2.5.5 Plant stress treatments (heat, salt, dehydration, cadmium chloride, SNP and sulfur starvation)	63
2.5.6 Pollen isolation, <i>in vitro</i> germination and FDA viability staining	65
2.5.7 Transcriptional profiling (microarray experiments)	66
2.5.7.1 Sample preparation, array hybridization and scanning	66
2.5.7.2 Data analysis	67
 Chapter 3: Alignments and Sequence Analysis	 69
3.1 Introduction	69
3.2 Results	70
3.2.1 The PQ-Loop domain	70
3.3.2 The PQ-loop repeat family in <i>A. thaliana</i>	73
3.3.2.1 <i>PQL</i> family sub-groupings	74
3.3.2.2 Sub-group 1: <i>AtPQL1</i> , <i>AtPQL2</i> , <i>AtPQL3</i> and <i>STM1</i>	75

3.3.2.3 Sub-group 2: <i>AtPQL4</i> , <i>AtPQL6</i> and <i>MPDU1</i>	78
3.3.2.4 Sub-group 3: <i>AtPQL5</i> and <i>CTNS</i>	78
3.3.3 Identification of potentially important residues in the PQ-loop	80
3.3.4 Similarity to non-PQL proteins	83
3.3.4.1 <i>AtPQL4</i> and <i>AtPQL6</i> and the 7TMR-DISM_7M domain	83
3.3.4.2 <i>AtPQL6</i> and the TDT domain	84
3.3.4.3 The <i>AtPQL</i> family and the TRP ion channel family	84
3.3.5 Genomic organisation of <i>AtPQL1-6</i>	88
3.3.6 Intron-exon organisation of <i>AtPQL1-6</i>	88
3.3.7 Promoter sequences of <i>AtPQL1-6</i>	90
3.4 Discussion	91
3.4.1 Functional relevance of the PQ-loop	91
3.4.2 Functional relevance of conserved topology	92
3.4.3 Intron-exon and genomic organisation	92
3.4.4 Specific features of individual genes and sub-groupings	93
3.4.4.1 Functional implications for sub-group 1: <i>AtPQL1-3</i>	93
3.4.4.2 Functional implications for sub-group 2: <i>AtPQL4</i> and <i>AtPQL6</i>	93
3.4.4.3 Functional implications for sub-group 3: <i>AtPQL5</i>	94
 Chapter 4: Expression Profile and Sub-cellular Localisation	 95
4.1 Introduction	95
4.1.1 Expression profile	95
4.1.2 Sub-cellular localisation	96
4.2 Results	97
4.2.1 Expression profile	97
4.2.1.1 Sub-group 1: <i>AtPQL1</i> , <i>AtPQL2</i> and <i>AtPQL3</i>	97
4.2.1.2 Sub-group 2: <i>AtPQL4</i> and <i>AtPQL6</i>	102
4.2.1.3 Sub-group 3: <i>AtPQL5</i>	105
4.2.2 Microarray databases	107
4.2.2.1 Microarray data - tissue-specific localisation	107
4.2.2.2 Microarray data - expression during pollen development	109
4.2.2.3 Microarray data - diurnal changes in expression	111
4.2.3 Sub-cellular localisation	112
4.2.3.1 Sub-group 1: <i>AtPQL1</i> , <i>AtPQL2</i> and <i>AtPQL3</i>	112
4.2.3.2 Sub-group 2: <i>AtPQL4</i> and <i>AtPQL6</i>	114
4.2.3.3 Sub-group 3: <i>AtPQL5</i>	114
4.3 Discussion	118
4.3.1 Unique features of individual family members and sub-groups	118
4.3.1.1 Sub-group 1: <i>AtPQL1</i> , <i>AtPQL2</i> and <i>AtPQL3</i>	118
4.3.1.2 Sub-group 2: <i>AtPQL4</i> and <i>AtPQL6</i>	119
4.3.1.3 Sub-group 3: <i>AtPQL5</i>	120
4.3.2 Common features amongst family members	121
4.3.2.1 Overlapping tissue-specific localisation	121
4.3.2.2 Response to viral infection and wounding	122
4.3.2.2 Diurnal changes in expression	123

Chapter 5: Functional Characterisation of the <i>AtPQL</i> Family	126
5.1 Introduction	126
5.1.1 Chapter organisation	126
5.1.2 Heterologous expression in yeast	126
5.1.3 Isolation and physiological characterisation of mutants	127
5.1.3.1 Isolation of mutant lines	127
5.1.3.2 Physiological characterisation of mutant lines	128
5.1.4 Transcriptional profiling of mutant lines	130
5.1.4.1 Microarray experiments	130
5.1.4.2 Links between glycosylation and gene expression	131
5.2 Results: Heterologous expression in yeast	131
5.2.1 Phenotype of <i>ers1Δ</i> deletion mutants	131
5.2.2 Expression of <i>AtPQL</i> proteins in yeast	135
5.3 Results: Isolation and physiological characterisation of mutants	137
5.3.1 Isolation of T-DNA insertion mutants	137
5.3.2 Genetic characteristics of T-DNA insertion mutants	139
5.3.3 Over-expresser mutant isolation	141
5.3.4 Phenotypic analysis of mutant lines	142
5.3.4.1 Analysis of growth in normal conditions	142
5.3.4.2 Sensitivity to heat shock	145
5.3.4.3 Sensitivity to salt: soil-grown plants	145
5.3.4.4 Sensitivity to salt: root-bending assay	147
5.3.4.5 Sensitivity to salt: seedling salt and osmotic stress tolerance assay	149
5.3.4.6 Sensitivity to dehydration	150
5.3.4.7 Analysis of leaf water loss	150
5.3.4.8 Sensitivity to cadmium	150
5.3.4.9 Sensitivity to SNP	152
5.3.4.10 Analysis of pollen viability and germination	155
5.4 Results: Transcriptional profiling of mutant lines	158
5.4.1 Microarray experiments	158
5.4.2 Altered gene expression in <i>AtPQL4</i> and <i>AtPQL6</i> mutants	159
5.4.3 Iterative group analysis (iGA)	161
5.4.4 Further analysis of differentially regulated genes	164
5.5 Discussion	166
5.5.1 Heterologous expression in yeast	166
5.5.2 Isolation and physiological characterisation of mutant lines	167
5.5.3 Transcriptional profiling of mutant lines	168
5.5.3.1 Defence and cell-surface related genes	169
5.5.3.2 Lipid-metabolism and ER stress-related genes	171
5.5.3.3 Secretory pathway-associated and N-glycoprotein encoding genes	173
5.5.3.4 Conclusions	174
Chapter 6: General Discussion	178
6.1 Assessment of <i>AtPQL</i> function in plants	178
6.1.1 Are any <i>AtPQL</i> proteins functionally homologous to STM1?	178

6.1.2 Are any <i>AtPQL</i> proteins functionally homologous to CTNS?	179
6.1.3 Are any <i>AtPQL</i> proteins functionally homologous to MPDU1?	179
6.1.4 Speculative role in proton translocation	180
6.1.5 Up-regulation by wounding and viral infection	181
6.1.6 Prospects for further study	183
Appendix A: Primer sequences	187
Appendix B: Supplementary microarray data	190
References	203

List of Figures and Tables

Figure 1-1. Structure of cysteine (cys) and cystine	2
Figure 1-2. LLO and GPI anchor biosynthesis in the ER	9
Figure 1-3. The calnexin-calreticulin cycle	13
Figure 2-1. Expression cassettes of Gateway expression vectors used in this study	46
Figure 2-2. Vector map of <i>pDR195 P_{PMAL}::AtPQL1-6</i> yeast expression clones	48
Figure 3-1. Hydropathy plots of selected PQ-loops	71
Figure 3-2. Sequence alignment of PQ-loops from various PQL proteins	72
Figure 3-3. Hydropathy plots of <i>AtPQL1-6</i>	73
Figure 3-4. Phylogenetic tree and domain organisation of selected PQL proteins.	74
Figure 3-5. Full-length alignment of <i>AtPQL1-6</i> , CTNS, ERS1, STM1 and MPDU1	76
Figure 3-6. Alignment of <i>AtPQL</i> proteins with their closest characterised non-plant homologues	79
Figure 3-7. Alignment of known functional residues in CTNS and STM1 with equivalents in <i>AtPQL1-6</i>	82
Figure 3-8. Similarity between the 7TMR-DISM 7TM domain and <i>AtPQL4</i> and <i>AtPQL6</i>	85
Figure 3-9. Similarity between <i>AtPQL6</i> and the TDT family	87
Figure 3-10. Alignment of consensus sequences for the <i>AtPQL</i> and TRP families	87
Figure 3-11. Intron-exon organisation of <i>AtPQL1-6</i>	89
Figure 4-1. Expression cassettes of plasmids used in this chapter	96
Figure 4-2. Expression level of <i>AtPQL1-3</i> as determined by real-time reverse-transcription PCR	99
Figure 4-3. Histochemical GUS staining of transgenic plants expressing an <i>AtPQL1</i> promoter::GUS construct	100
Figure 4-4. Transcript level of <i>AtPQL4</i> and <i>AtPQL6</i> as determined by real-time reverse-transcription PCR	103
Figure 4-5. Transcript level of <i>AtPQL5</i> as determined by real time reverse transcription PCR	106
Figure 4-6. Microarray expression data for <i>AtPQL1</i> , <i>AtPQL2</i> , <i>AtPQL5</i> and <i>AtPQL6</i> in different tissues	108
Figure 4-7. Microarray expression data for <i>AtPQL1</i> , <i>AtPQL2</i> , <i>AtPQL5</i> and <i>AtPQL6</i> during pollen development	110

Figure 4-8. Microarray expression data for <i>AtPQL5</i> and <i>AtPQL6</i> over the course of two days	111
Figure 4-9. Sub-cellular localisation of <i>AtPQL1</i>	113
Figure 4-10. Sub-cellular localisation of <i>AtPQL4</i>	115
Figure 4-11. Sub-cellular localisation of <i>AtPQL6</i>	116
Figure 4-12. Sub-cellular localisation of <i>AtPQL5</i>	117
Figure 4-13. Relative tissue-specific expression level of <i>AtPQL1-6</i>	122
Figure 4-14. Microarray expression data for <i>AtPQL6</i> , and <i>ALG1/ALG2</i> - and <i>GPT</i> -like genes over the course of two days	125
Figure 5-1. Growth of yeast strains on hygromycin	133
Figure 5-2. Growth of yeast strains under a variety of conditions	134
Figure 5-3. Growth of yeast strains on low and high glucose media	136
Figure 5-4. Growth of yeast strains on cystine-containing media	137
Figure 5-5. T-DNA insertion sites of SALK mutant lines used in this study	137
Figure 5-6. Confirmation of T-DNA knockout lines at the gDNA level	138
Figure 5-7. Confirmation of T-DNA knockout lines at the transcript level	140
Figure 5-8. Expression cassettes of plasmids used in this chapter	140
Figure 5-9. Confirmation of <i>AtPQL</i> over-expression	143
Figure 5-10. Images of wild type and <i>AtPQL6</i> mutant lines	144
Figure 5-11. Heat shock of mutant lines with altered expression of <i>AtPQL4</i> and <i>AtPQL6</i>	146
Figure 5-12. Salt stress treatment of mutant lines with altered expression of <i>AtPQL4</i> and <i>AtPQL6</i>	147
Figure 5-13. Root-bending assay of <i>pql4-1</i> salt tolerance	148
Figure 5-14. Seedling salt and osmotic stress tolerance assay	149
Figure 5-15. Dehydration of mutant lines with altered expression of <i>AtPQL4</i> and <i>AtPQL6</i>	151
Figure 5-16. Water loss from excised leaves of wild-type and <i>pql5-1</i>	152
Figure 5-17. Cadmium tolerance of <i>AtPQL</i> mutants in control and low sulfur conditions	153
Figure 5-18. Effect of the nitric oxide donor SNP on <i>AtPQL</i> mutant lines in control and low sulfur conditions	154
Figure 5-19. Viability of <i>AtPQL</i> mutant pollen	156
Figure 5-20. <i>In vitro</i> germination of <i>AtPQL</i> mutant pollen	157
Figure 5-21. Genes showing common or opposite regulation in <i>AtPQL</i> mutants	160
Figure 5-22. Differential regulation of <i>BiP3</i> in <i>AtPQL4</i> and <i>AtPQL6</i> mutants	161
Figure 5-23. Proposed explanation for molecular phenotype of <i>AtPQL4</i> and <i>AtPQL6</i> mutants	177
Figure B-1. Sammon mapping of microarray data	190
Table 1-1. Summary of mutants affected in Man/Glc-P-Dol dependent glycosylation in <i>A. thaliana</i>	21
Table 2-1. Reagent concentration in control and low sulfur medium	38
Table 2-2. Standard reaction mix and cycling conditions for PCR	40
Table 2-3. Standard reaction mix and cycling conditions for real-time reverse-transcription PCR	42
Table 2-4. Plasmids produced in this study by gateway and conventional cloning	45

Table 2-5. Affymetrix ATH1 probe identifiers for <i>AtPQL</i> , <i>ALG1/2-like</i> and <i>GPT-like</i> genes	58
Table 2-6. Yeast media recipes	59
Table 3-1. Names and accession numbers of PQL proteins referred to in this chapter	71
Table 3-2. Pairwise comparisons of PQL proteins	77
Table 3-3. Point mutations in CTNS and MPDU1	82
Table 3-4. Blast searches using PQL proteins as queries	86
Table 3-5. Homologous <i>PQL</i> genes in duplicated regions of the <i>A. thaliana</i> and <i>P. trichocarpa</i> genomes	89
Table 3-6. Genetic characteristics of <i>AtPQL1-6</i>	89
Table 3-7. Known transcription factor motifs upstream of <i>AtPQL1-6</i>	90
Table 5-1. Yeast strains used in this study	132
Table 5-2. Segregation analysis of <i>AtPQL</i> over-expresser lines	142
Table 5-3. Zygoty testing of the progeny of heterozygous <i>pql4</i> and <i>pql6</i> mutants	156
Table 5-4. Number of differentially regulated genes in <i>AtPQL</i> mutants	159
Table 5-5. Groups of functionally related genes showing differential regulation in <i>AtPQL4</i> mutants	162
Table 5-6. Groups of functionally related genes showing differential regulation in <i>AtPQL6</i> mutants	163
Table 5-7. Predicted sub-cellular localisation and N-glycosylation status of proteins encoded by differentially regulated genes in <i>AtPQL4</i> and <i>AtPQL6</i> mutants	165
Table 6-1. Summary of proposed localisation, function and regulation of <i>AtPQL</i> proteins	182
Table A-1. Primer sequences for qPCR	187
Table A-2. Primer sequences for Gateway pENTR-D-TOPO cloning	187
Table A-3. Primer sequences for Gateway BP cloning	188
Table A-4. Primers for cloning <i>AtPQL</i> open reading frames into pDR195 vectors	188
Table A-5. Primers for T-DNA insertion line zygoty testing	189
Table A-6. Other primers used in this study	189
Table B-1. Genes up-regulated in <i>pql4-1</i> relative to wild type	191
Table B-2. Genes up-regulated in <i>pql4-2</i> relative to wild type	191
Table B-3. Genes down-regulated in <i>pql4-2</i> relative to wild type	192
Table B-4. Genes down-regulated in <i>pql6-1</i> relative to wild type	192
Table B-5. Genes up-regulated in <i>pql6-1</i> relative to wild type	193
Table B-6. Genes up-regulated in <i>35S::AtPQL6-1</i> relative to wild type	195
Table B-7. Genes down-regulated in <i>35S::AtPQL6-1</i> relative to wild type	202

Terminology

The terms “up-regulated” and “down-regulated” are frequently used in chapter 5 when referring to genes with altered transcript level between mutant and wild type (as measured by microarrays). In this context regulation can reflect changes in both gene transcription and mRNA turnover.

Abbreviations

ABA	abscisic acid	iGA	iterative group analysis
ATP	adenosine triphosphate	JA	jasmonic acid
cAMP	cyclic adenosine monophosphate	LCT	lysosomal cystine transporter
CaMV	cauliflower mosaic virus	LLO	lipid-linked oligosaccharide
CDG-If	congenital disorder of glycosylation If	Man	mannose
cDNA	complimentary DNA	Man-1-P	mannose-1-phosphate
CHO	Chinese hamster ovary	Man-6-P	mannose-6-phosphate
Cit	citronellol	Man-P-Cit	mannose-phosphate-citronellol
ConA	concanavalin A	Man-P-Dol	mannose-phosphate-dolichol
Cs	castanospermine	MPDU1	Man-P-Dol utilisation defect 1
CTNS	cystinosis	MR	microbial rhodopsins
Cys	cysteine	mRNA	messenger RNA
Dedol-P-P	dehydrodolichol pyrophosphate	MS	Murashige and Skoog
DNA	deoxyribonucleic acid	NO	nitric oxide
dNTP	deoxyribonucleotide triphosphate	OD	optical density
Dol	dolichol	ORF	open reading frame
Dol-P	dolichol monophosphate	OST	oligosaccharyltransferase
Dol-P-P	dolichol pyrophosphate	PCR	polymerase chain reaction
ER	endoplasmic reticulum	PGM	pollen germination medium
ERS1	<i>erd1A</i> suppressor 1	PI	phosphatidylinositol
EtNP	ethanolaminephosphate	PMM	phosphomannomutase
FDA	fluorescein diacetate	PQL	PQ-loop repeat
gDNA	genomic DNA	qPCR	quantitative PCR
GDP	guanosine diphosphate	RGS	regulator of G-protein signalling
GDP-Man	guanosine diphosphate-mannose	RNA	ribonucleic acid
GEF	guanine nucleotide exchange factor	RP	rank product
GFP	green fluorescent protein	SA	salicylic acid
Glc	glucose	SAR	systemic acquired resistance
GlcNAc	N-acetylglucosamine	SDS	sodium dodecyl sulphate
Glc-P-Cit	glucose-phosphate-citronellol	SLO	streptolysin-O
Glc-P-Dol	glucose-phosphate-dolichol	SNP	sodium nitroprusside
GNSO	S-nitrosoglutathione	STM1	seven transmembrane 1
GPCR	G-protein coupled receptor	Sw	swainsonine
GPI	glycophosphatidylinositol	TAE	tris acetyl edta
G-protein	guanine nucleotide binding protein	T-DNA	transferred DNA
GPT	GlcNAc phosphotransferase	UDP	uridine diphosphate
GSH	glutathione	UDP-Glc	uridine diphosphate-glucose
GT	UDP-Glc:glycoprotein glucosyltransferase	UDP-GlcNAc	uridine diphosphate-N-acetylglucosamine
GTP	Guanosine triphosphate	UTR	untranslated region
HPLC	high performance liquid chromatography	UV	ultraviolet
		YPD	yeast peptone dextrose medium

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Anna Amtmann for the opportunity to work on this project and for continuing to support me throughout even as the project led in unexpected directions. I greatly appreciate the advice and guidance she has provided me during my time in the lab and the encouragement she has given me to succeed.

I would also like to thank past and present members of the Blatt/Amtmann labs for their help and friendship over the past three years. In particular, Patrick Armengaud whose guidance in general laboratory practice and advice on technical matters has been invaluable. I also thank Ufo Sutter, Manuel Paneque and George Littlejohn for help with confocal microscopy, cloning and plant transformation and George Boswell and Naomi Donald for help with seed collection. In addition I thank Pawel Herzyk and the SHWFG facility for microarray hybridisation and analysis.

Finally, I thank Stacey for support throughout and my parents for giving me direction and having confidence in my ability.

Abstract

This thesis describes the characterisation of the PQ-loop repeat family, a novel family of membrane proteins from *Arabidopsis thaliana*. The family consists of six members (*AtPQL1-6*) all predicted to contain seven transmembrane segments and two copies of the PQ-loop domain. The repeated PQ-loop domain is a well conserved feature present in several eukaryotic proteins. However, the functional significance of the domain remains to be determined. Although most PQ-loop repeat proteins are of unknown function some have been characterised from non-plant species. Those studied include CTNS - a lysosomal cystine transporter from humans, STM1 - a putative G-protein coupled receptor from *Schizosaccharomyces pombe*, ERS1 - a *Saccharomyces cerevisiae* protein isolated as a suppressor of a trafficking mutant and LEC35/MPDU1 - an endoplasmic reticulum protein from mammals involved in the utilisation of dolichol monosaccharides during glycosylation.

The *AtPQL* family can be divided into three groups based on similarity with their non-plant homologues. Detailed sequence analysis was carried out to investigate conserved motifs, topology, domain organisation and genome distribution of the six genes. The implications of these findings for putative functions are discussed.

AtPQL gene expression was investigated using promoter::GUS fusions and quantitative PCR. Whereas *AtPQL4* and *AtPQL6* appear to be expressed ubiquitously, *AtPQL1*, *AtPQL2* and *AtPQL5* are all preferentially expressed in floral tissues and *AtPQL3* is expressed primarily in roots. Several family members showed rhythmic expression over the course of a day and induction by both viral infection and wounding. Localisation at the sub-cellular level was investigated using green fluorescent protein fusions and confocal microscopy. At least four *AtPQL* proteins are targeted to intracellular compartments. *AtPQL1* and *AtPQL5*-GFP fusions both highlight the tonoplast whereas *AtPQL4* and *AtPQL6*-GFP appear to highlight the ER.

A number of knockout and over-expresser mutants were isolated for each gene. Although no physiological phenotype could be found, mutants with altered expression of *AtPQL4* and *AtPQL6*, the two closest *A. thaliana* homologues of *MPDU1/LEC35*, showed a clear molecular phenotype. These mutants displayed differential regulation of a number of genes encoding N-glycoproteins as well as defence, cell-surface, lipid biosynthesis and ER stress associated proteins.

Finally, the functional importance of the *AtPQL* family is discussed and a role for *AtPQL4* and *AtPQL6* in glycosylation proposed.

Chapter 1: Introduction

This thesis details the sequence analysis, expression profiling and functional characterisation of a novel protein family in *Arabidopsis thaliana*. The family, named PQ-loop repeat (or PQL), consists of six members, *AtPQL1-6*, none of which have been characterised previously.

Family members are defined by the presence of two PQ-loop domains. The PQ-loop is a conserved region of 40-60 amino acids named after a highly conserved proline-glutamine (PQ) dipeptide motif. *AtPQL* family members are all integral membrane proteins and each is predicted to have seven transmembrane spanning domains.

A detailed sequence analysis of the *AtPQL* family is provided in chapter 3 along with an in-depth description of family characteristics. In this chapter I will review the existing literature on PQL proteins, the vast majority of which have not been characterised. However, a small number have been studied in non-plant species and play apparently diverse roles in cellular processes including amino acid homeostasis (Kalatzis et al., 2001), glycosylation (Hardwick and Pelham, 1990; Anand et al., 2001) and nutrient sensing (Chung et al., 2001). For these PQL proteins I will discuss their identification, characterisation and functional significance in their own organism. I will also discuss the potential relevance of these proteins should their homologues play similar roles in plants.

1.2 Cystinosin (CTNS)

The first PQL protein I will discuss is cystinosin (CTNS), a human lysosomal cystine/proton symporter (Kalatzis et al., 2001). Mutations in *CTNS* are responsible for cystinosis; a disease where cystine hyper-accumulates in lysosomes (Town et al., 1998). Hyper-accumulation results in the build up of insoluble cystine crystals in almost all cells and tissues and severe mutations lead to poor growth, renal failure, hypothyroidism and photophobia (Gahl et al., 2002). CTNS is unique amongst PQL proteins in that it is the only family member for which both the precise biological role and molecular mode of action are known.

Cystine, the disulfide form of the common protein forming amino acid and key sulfur metabolite cysteine (cys; see figure 1-1), is produced in lysosomes as a by-product of protein hydrolysis (Gahl et al., 2002). Under normal conditions, cystine is exported to the cytosol and reduced to cys, but during cystinosis accumulates in the lysosomes where oxidizing conditions prevent reduction (Gahl et al., 2002).

CTNS was identified as the gene responsible for cystinosis by map-based cloning and shown to encode an integral membrane protein (Town et al., 1998). Sub-cellular localisation studies identified a C-terminal signalling motif (GYDQL) critical for lysosomal targeting (Cherqui et al., 2001). Deletion mutants lacking the GYDQL targeting motif were utilised for functional characterisation (Kalatzis et al., 2001). When expressed in COS cells, the mutant forms localised to the plasma membrane and as a result, the intra-lysosomal surface of the protein faced the external medium. This allowed transport activity to be assayed via radiolabelled cystine uptake. Using this system it was demonstrated that CTNS was able to transport cystine. It was further reported that CTNS-mediated cystine transport was proton coupled as transport could be stimulated by acidification of the external medium and inhibited by the use of nigericin, an ionophore used to disrupt proton gradients.

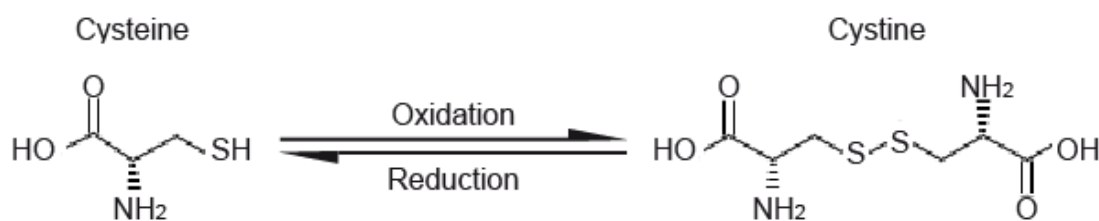


Figure 1-1. Skeletal models of cysteine (cys; C00097) and cystine (C00491). Image adapted from KEGG (Kanehisa and Goto, 2000). The two forms are readily interchangeable by non-enzymatic oxidation and reduction reactions.

1.2.1 Biological importance of cystine

In contrast to the problems caused by cystine hyper-accumulation, cystine can play positive roles in cellular processes. Firstly, cystine can be reduced to cys. The two forms are readily interchangeable but under normal physiological conditions, cys is predominant (Cooper, 1983). In plants, cys is critical in sulfur nutrition, representing the end point of sulfur assimilation and a key intermediate in sulfur metabolism (Saito, 2004). As well as incorporation into proteins, cys is used for the biosynthesis of a variety of sulfur containing compounds such as glutathione (GSH), phytochelatins, methionine and glucosinolates. The major source of cellular cys is the cys synthase complex, a hetero-oligomer consisting of serine acetyltransferase and O-acetylserine thiol lyase subunits which catalyze the production of cys from sulfide, serine and acetyl CoA (Wirtz and Hell, 2006). Relatively little is known about the degradation and recycling of cys in plants (Saito, 2004). However, in animals cystine exported from lysosomes does contribute to overall cys availability, and as a result, cystinotic cell lines show decreased levels of GSH (Chol et al., 2004).

In addition, cystine itself is the preferred substrate of C-S lyases, a class of enzymes that cleave the β -carbon-sulfide link of L-cystine (or other disulfides) to produce thiocysteine, ammonia and pyruvate (Hamamoto and Mazelis, 1986). The thiocysteine produced in this reaction is unstable and rapidly reacts with free cystine to produce thiocystine which can then act as a sulfur donor for rhodanases - a class of enzymes that detoxify cyanide to the less toxic, thiocyanate (Pinto et al., 2006). C-S lyases have been purified from broccoli (Hamamoto and Mazelis, 1986) and identified at the molecular level in *A. thaliana* (Jones et al., 2003; Mikkelsen et al., 2004). However, these enzymes do not exclusively use cystine as a substrate. One C-S lyase from *A. thaliana*, *SURI*, cleaves the β -carbon-sulfide link of S-alkylthiohydroximate producing thiohydroximate for glucosinolate biosynthesis (Mikkelsen et al., 2004).

1.2.2 CTNS-like function in plants

As mentioned above, in animals the main source of lysosomal cystine is protein hydrolysis (Gahl et al., 2002). In plants the majority of proteinases are localised to the vacuole (Callis, 1995), so a functional homologue of CTNS would presumably export cystine from vacuoles. As of yet no vacuolar cystine transporters have been identified in plants and it is unclear whether plants would require such a protein. An obvious difference between vacuoles and lysosomes is size. Lysosomes are small organelles, typically 1 μm in diameter (Toyomura et al., 2003) whereas plant vacuoles are large structures often consisting of up to 90% of the cell volume (Alberts, 2002). Whereas lysosomes may require cystine export to prevent over-accumulation, vacuoles may be large enough to store all the cystine produced by protein hydrolysis without adverse effects. Another difference, is that whereas lysosomes specifically function in degradation and maintain a low pH and an oxidizing environment, plant vacuoles have more diverse functions such as sequestration of toxins, osmoregulation and storage of ions and macromolecules (Alberts, 2002). It is unclear whether the highly oxidising conditions preventing cystine reduction in mammalian lysosomes are as severe in plant vacuoles.

If CTNS-like function is required in plants and *AtPQL5*, the plant PQL protein with the closest sequence homology to CTNS, is functionally homologous, a mutant plant lacking this protein may be expected to show a number of phenotypes. Firstly, this mutant may show severe growth defects as a result of cystine hyper-accumulation and crystallisation as seen during cystinosis (Gahl et al., 2002). In addition, a CTNS-like mutant may have decreased levels of cytoplasmic cystine resulting in a lack of cystine as a substrate for C-S lyases and, as a knock-on effect, shortages of the reduced form, cys. As cys is critical for the biosynthesis of sulfur-containing compounds, such a mutant may show sulfur-deficiency symptoms, typically reduced growth, decreased shoot-root biomass ratio and leaf chlorosis (Marschner, 1995). There might also be stress related phenotypes since several sulfur-containing compounds (e.g. GSH and phytochelatins) play important roles in stress responses (Saito, 2004).

It has been reported that exogenous application of cystine to *A. thaliana* causes increased tolerance to cadmium, presumably due to increased synthesis of GSH and phytochelatins (Dominguez-Solis et al., 2001). Therefore it may be expected that over-expresser mutants of a CTNS-like protein in plants would show a similar phenotype. Although, cytoplasmic cys availability may not be significantly reduced in a CTNS-like knockout, it may be that cytoplasmic cys is increased in an over-expresser leading to increased stress tolerance as a result of increased GSH and phytochelatin production.

1.3 MPDU1 and ERS1

Two PQL proteins, MPDU1/LEC35 from mammals and ERS1 from *S. cerevisiae*, have been implicated in glycosylation. MPDU1, originally identified as LEC35 in Chinese hamster ovary (CHO) cells is required for the use of mannose-phosphate-dolichol (Man-P-Dol) and glucose-phosphate-dolichol (Glc-P-Dol) as sugar donors for glycosyltransferases (Anand et al., 2001). Two *AtPQL* proteins which are very similar to each other, *AtPQL4* and *AtPQL6* are the closest *A. thaliana* homologues of MPDU1 (see chapter 3). The role of ERS1 is unclear, but was isolated as a suppressor of a yeast mutant with defects in glycosylation and protein trafficking (Hardwick and Pelham, 1990).

1.3.1 Man-P-Dol and Glc-P-Dol dependent glycosylation

Glycosylation is the covalent linkage of sugar residues to proteins. It is an important form of post-translational modification that is well conserved throughout eukaryotes (Lehle et al., 2006). Five distinct classes of amino acid-sugar linkages have been reported, each with its own specific biosynthetic pathway and sugar donor requirements (Spiro, 2002).

The most common class, N-glycosylation is dependent on both Man-P-Dol and Glc-P-Dol as sugar donors (Helenius and Aebi, 2004). C-mannosylation and GPI anchoring require Man-P-Dol but not Glc-P-Dol (Doucey et al., 1998; Kinoshita and Inoue, 2000) whereas O-glycosylation is more heterogeneous and may use Man-P-Dol

depending on the specific oligosaccharide chain (Strahl-Bolsinger et al., 1999). The biosynthetic pathway for P-glycosylation is less clear, but it appears that neither Man-P-Dol or Glc-P-Dol are required (Moss et al., 1999).

1.3.1.1 Synthesis, reorientation and utilisation of Man-P-Dol and Glc-P-Dol

Depending on the type of glycosylation, sugar residues are linked either directly to a protein or to a lipid-linked precursor. These linkage reactions are catalysed by glycosyltransferases localised to either the cytosolic or luminal sides of the ER and golgi membranes. In general, those taking place on the cytosolic side use nucleotide sugars (UDP-Glc, UDP-GlcNAc, UDP-Man) as sugar donors (Schenk et al., 2001a). However, unlike in the golgi, where nucleotide sugars are transported across the membrane (Handford et al., 2004), nucleotide sugars are not usually used as sugar donors in the ER lumen. Instead the sugar residues are linked to dolichol monophosphate (Dol-P), which acts a carrier molecule, allowing the sugar residue to be flipped across the membrane to face the luminal side (Schenk et al., 2001a).

The cytoplasmic linkage of Man and Glc to Dol-P is catalysed by Dol-P-mannosyltransferase and Dol-P glucosyltransferase using GDP-Man and UDP-Glc as substrates (see figure 1-2, steps 9 and 15; Helenius et al. 2001). Although their molecular identity has not yet been established in plants, enzyme activity has been demonstrated (Brett and Leloir, 1977; Chadwick and Northcote, 1980).

In order to be used as sugar donors, Man-P-Dol and Glc-P-Dol need to be re-orientated so that the sugar residue faces the luminal side of the membrane (Schenk et al., 2001a). This process has been studied using Man-P-Dol and Glc-P-Dol analogues, Man-P-citronellol (Man-P-Cit) and Glc-P-citronellol (Glc-P-Cit), which are structurally similar but whereas Dol has 18-19 isoprene units in the lipid chain, Cit has just two and is water soluble (Rush and Waechter, 1995, 1998). It was found that uptake of radiolabelled Man-P-Cit and Glc-P-Cit into isolated rat-liver vesicles was protein-dependent suggesting the involvement of a flippase. There was no requirement for ATP and it appears that the flippases for Man-P-Cit and Glc-P-Cit are independent as neither substrate inhibited uptake of the other.

The molecular identity of the Man-P-Dol and Glc-P-Dol flippases is unknown. It has been suggested that LEC35/MPDU1 may be the missing flippase, but there are also arguments favouring an alternate role (see section 1.3.7; Anand et al., 2001).

1.3.1.2 Mutations in A. thaliana affecting Man-P-Dol synthesis

In *A. thaliana*, mutations have been isolated that impair synthesis of Man-P-Dol by preventing synthesis of the Man-P-Dol precursor, GDP-Man. Two allelic mutants, *vtc1* and *cyt1*, have been described that affect mannose-1-phosphate (Man-1-P) guanylyltransferase, the enzyme responsible for the production of GDP-Man from Man-1-P and GTP (Conklin et al., 1996; Lukowitz et al., 2001). As well as being the Man donor for Man-P-Dol synthesis, GDP-Man is the Man donor for the first five Man residues in N-glycan biosynthesis (see figure 1-2, steps 4-7; Helenius and Aebi, 2004), and is also required for the production of ascorbic acid (Conklin et al., 1999). Furthermore, the presence of multiple golgi-localised proteins capable of GDP-Man transport suggests utilisation of GDP-Man in the golgi lumen perhaps for the modification of glycosylceramides and glycosylinositol phosphoryl ceramides (Handford et al., 2004).

cyt1 mutants have an embryo lethal phenotype with abnormalities appearing at the heart stage of embryo development and defects in cell division leading to enlarged and disorganized cells (Nickle and Meinke, 1998). Defects in N-glycosylation of protein disulfide isomerase has also been shown (Lukowitz et al., 2001). The *vtc1* mutation is not completely null and gives a much milder phenotype, not leading to lethality, but causing reductions in ascorbic acid biosynthesis (Conklin et al., 1999).

Recently, a temperature-sensitive point mutation has been isolated in *A. thaliana* affecting phosphomannomutase (PMM), the enzyme responsible for the conversion of Man-6-P to Man-1-P (Hoeberichts et al., 2008). As a result the mutants cannot produce GDP-Man (PMM is one reaction upstream of CYT1 in GDP-Man biosynthesis). The mutant, *pmm-12*, is lethal at the restrictive temperature and shows deficiencies in both N-glycosylation and GPI anchoring. A second mutant allele containing a T-DNA insertion 5 bp upstream of the start codon, gave an embryo lethal phenotype similar to *cyt1*.

1.3.2 N-linked glycosylation

I will now describe the four different classes of glycosylation known to require Man/Glc-P-Dol, paying particular attention to plants. Firstly, I will discuss N-linked glycosylation, so called because it involves the covalent linkage of an oligosaccharide (N-glycan) to an asparagine (N) residue in nascent proteins (Helenius and Aebi, 2004).

It should be noted that the term N-glycosylation is occasionally used to refer to the covalent linkage of sugars (usually Glc) to the amino (N) terminal of xenobiotics (Lao et al., 2003; Brazier-Hicks et al., 2007). This is a distinct process, most commonly using UDP-Glc as a sugar donor (Brazier-Hicks et al., 2007). In this thesis, N-glycosylation refers to asparagine-linked as opposed to amino group-linked glycosylation.

1.3.2.1 Lipid-linked oligosaccharide synthesis

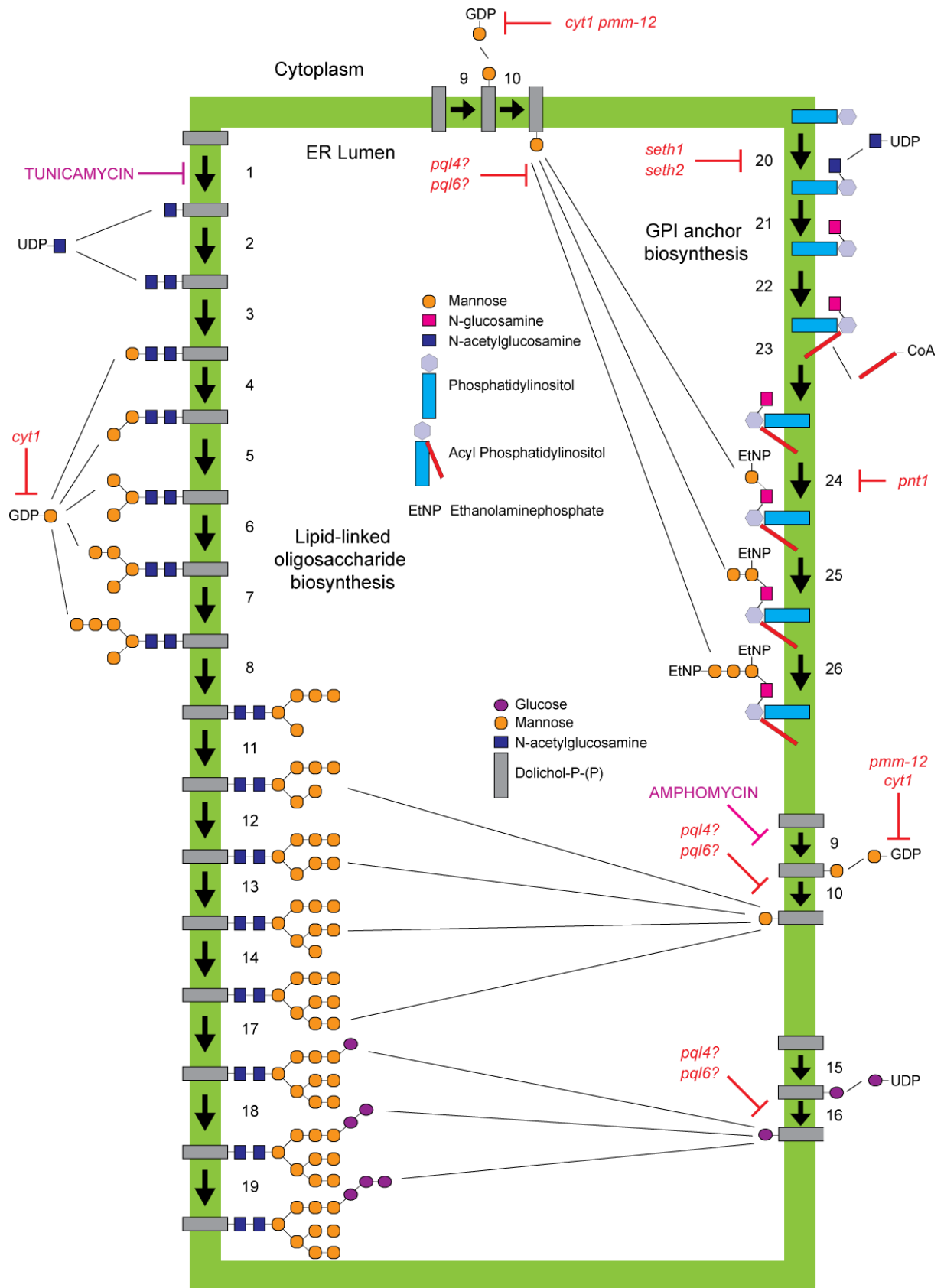
N-glycosylation begins with the formation of a lipid-linked oligosaccharide (LLO) precursor (Helenius and Aebi, 2004). Biosynthesis occurs on both sides of the ER membrane by addition of sugar residues to dolichol-pyrophosphate (Dol-P-P) by monosaccharyltransferases (see figure 1-2, steps 1-19).

First, two N-acetylglucosamine (GlcNAc) residues and five mannose (Man) residues are added to Dol-P-P on the cytoplasmic side of the ER membrane to produce $\text{Man}_5\text{GlcNAc}_2\text{-P-P-Dol}$ (see figure 1-2, steps 1-7) (Helenius and Aebi, 2004). In these reactions the nucleotide sugars, UDP-GlcNAc and GDP-Man, are used as sugar donors.

The LLO, $\text{Man}_5\text{GlcNAc}_2\text{-P-P-Dol}$, is then flipped by an ATP-independent flippase to face the luminal side of the ER (Helenius et al., 2002). Four additional Man and three Glc residues are then added to form $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-P-P-Dol}$ (see figure 1-2, steps 16-19). The sugar donors for these reactions are Man-P-Dol and Glc-P-Dol.

The oligosaccharide is then transferred from lipid to protein by the oligosaccharyltransferase (OST) complex (see section 1.3.2.3) and further modified by removal of three Glc and one Man residue as part of the calnexin-calreticulin pathway; an important quality control mechanism ensuring correct protein folding before export to the golgi for further processing (see section 1.3.2.5 ; Helenius and Aebi 2004).

Figure 1-2. LLO and GPI anchor biosynthesis in the ER. Lipid linked oligosaccharide synthesis for N-glycosylation (left) begins on the cytoplasmic side of the ER membrane with the transfer of two GlcNAc and five Man residues from nucleotide sugars to Dol-P-P forming $\text{Man}_5\text{GlcNAc}_2\text{-P-P-Dol}$ (**1-7**). The LLO is then flipped to face the luminal side (**8**) and four more Man (**11-14**) and three Glc (**17-19**) residues are added to form $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-P-P-Dol}$. The sugar donors for these reactions are Man-P-Dol and Glc-P-Dol, which are formed on the cytoplasmic side of the ER membrane by the transfer of Man and Glc from nucleotide sugars to Dol-P (**9,15**) before being flipped to face the lumen (**10,16**). GPI anchor biosynthesis (right) begins with the addition of GlcNAc to PI (**20**), deacylation of GlcNAc (**21**) and acylation of PI (**22**) on the cytoplasmic side of the ER membrane before the precursor is flipped to face the luminal side (**22**) where three Man residues are added from Man-P-Dol (**23-25**). The Man residues are modified by the addition of EtNP for linkage to target proteins. *A. thaliana* mutants affected in these processes are shown in red. Inhibitors are shown in purple.



1.3.2.2 Chemical inhibitors of LLO synthesis

Although several genes involved in N-glycan modification in the golgi have been studied extensively in plants (Strasser et al., 2007), relatively few have been identified with direct involvement in LLO synthesis. No loss-of-function mutants specifically affecting LLO synthesis have been described. However, work with tunicamycin, a specific inhibitor of GlcNAc phosphotransferase (GPT), the enzyme which catalyses the first glycosyltransferase reaction in LLO synthesis (see figure 1-2, step 1) gives some indication of the importance of this process in plants.

Treatment of cell cultures with tunicamycin is lethal for all eukaryotic cell types including plants (Zeng and Elbein, 1995). An early response to sub-lethal doses of tunicamycin is induction of endoplasmic reticulum stress related genes associated with the unfolded protein response (UPR), including signalling components and molecular chaperones (Urade, 2007). There are also indications of altered activity of key enzymes involved in lipid metabolism (Shank et al., 2001).

N-glycoproteins immuno-precipitated from plant tissue treated with tunicamycin are incorrectly folded and remain attached to BiP, a key ER localised molecular chaperone and protein folding determinant (D'Amico et al., 1992). Effects on trafficking have also been reported with tunicamycin preventing the secretion of many extracellular N-glycoproteins (Faye and Chrispeels, 1989) and causing vacuolar N-glycoproteins to be retained in the ER (Sparvoli et al., 2000).

Further effects of tunicamycin have been described suggesting a link to programmed cell death and systemic acquired resistance, both of which are important pathogen defence mechanisms. Treatment of *Cucumis sativus* (cucumber) with tunicamycin and amphomycin (an inhibitor of Man-P-Dol synthase) led to increased systemic acquired resistance (SAR) against a fungal pathogen, *Colletotrichum lagenarium*, with tunicamycin-treated plants showing increased expression of PR proteins, increased levels of salicylic acid (SA) and cell death at the point of fungal application (Sticher and Metraux, 2000). In addition prevention of N-glycosylation with tunicamycin in cultured tobacco cells led to cell death and induction of a marker gene, *HSR203J*, for programmed cell death as part of the hypersensitive response (Iwata and

Koizumi, 2005a). However, the mechanisms linking N-glycosylation and defence responses are unknown.

In several species tunicamycin resistance can be conferred by over-expressing GPT (Lehrman et al., 1988). In plants, a tunicamycin-resistant soybean cell line has been isolated by step-wise selection and shown to have increased GPT activity (Zeng and Elbein, 1995). Similarly in *A. thaliana*, where *GPT* has been cloned, an over-expresser mutant showed both increased tolerance to tunicamycin and reduced expression of *BiP* (Koizumi et al., 1999).

1.3.2.3 N-Glycan oligosaccharyltransferase mutants

Although no loss-of-function mutants have been described which specifically affect LLO synthesis, several mutants have been reported that affect components of the oligosaccharyltransferase (OST) complex responsible for transferring the oligosaccharide to proteins.

The first OST loss-of-function mutants described in *A. thaliana* were *stt3a* and *stt3b* (Koiwa et al., 2003). *stt3a* was first identified in a screen for mutants hypersensitive to salt stress. In addition, *stt3a* had reduced expression of a meiosis specific marker gene and partly reduced glycosylation of known N-glycoproteins. Sequence analysis showed homology to components of the yeast OST and another *A. thaliana* gene named *SST3B*. *sst3b* single mutants show no obvious phenotype but *sst3a sst3b* double mutants show gametophytic lethality through both pollen and ovules.

Another OST mutant, *dgl1-1*, was identified in a screen for mutants with hypocotyl growth defects and altered composition of cell wall polysaccharides (Lerouxel et al., 2005). Despite being leaky and still producing transcript, *dgl1-1* mutants stopped growing at an early stage of seedling development, showed reduced glycosylation of known glycoproteins and had multiple cell wall defects. A second (presumably null) allele, *dgl1-2* was embryo lethal.

AtDAD1 and *AtDAD2*, the anchoring components of the OST, have also been cloned and shown to protect against UV induced DNA fragmentation and programmed cell death when over-expressed in protoplasts (Danon et al., 2004). However, this

function was independent of inhibition of glycosylation inhibition by tunicamycin suggesting the proteins have pleiotropic and distinct functions in OST anchoring and programmed cell death. *dad1* knockout mutants were unable to secrete PR-1, a known glycoprotein, in response to SA treatment (Wang et al., 2005).

1.3.2.4 N-Glycosylation and the calnexin-calreticulin cycle

Other mutants have been described that act shortly after the OST complex but before export to the golgi. They affect glucosidase-I (Boisson et al., 2001; Gillmor et al., 2002) and glucosidase-II (Taylor et al., 2000; Burn et al., 2002). These enzymes trim terminal Glc residues from N-glycans and play an important role in quality control ensuring correct protein folding through the calnexin-calreticulin cycle (see figure 1-3; Helenius and Aebi 2001).

After glucosidase I and II trim terminal Glc residues, one Glc may be re-added to the glycan by UDP-Glc:glycoprotein glucosyltransferase (GT; Helenius and Aebi 2001). GT specifically recognises and re-glucosylates only incompletely folded proteins. After re-glucosylation by GT, the glycan is bound by calnexin and calreticulin, two lectins that prevent export from the ER and promote correct protein folding. Once correct folding has been established the protein is no longer re-glucosylated after Glc trimming by glucosidase II, no longer bound by calnexin and calreticulin, and export to the golgi can occur.

The mutations affecting glucosidase I, *gcs1* (Boisson et al., 2001) and *knf1* (Gillmor et al., 2002) both cause shrunken, wrinkled seeds and deficiencies in embryo development. Seeds of both mutants were unable to germinate and both showed reduced transmission through pollen. *gcs1* also showed reduced transmission through ovules. A glucosidase II mutant, *rsw3*, has a temperature sensitive radial swelling phenotype (Burn et al., 2002). *rsw3* mutants also showed a defective seed set and reduced male and female transmission. A second glucosidase II mutant, an antisense mutant from potato, showed a strong growth retarded phenotype in field conditions and increased expression of *BiP*, a strong indicator of ER stress (Taylor et al., 2000).

Another mutant has been described which lacks an ER localised UDP-Glc transporter, *AtUTr1* (Reyes et al., 2006). Unlike the other glucosylation reactions in the ER lumen which use Glc-P-Dol as a Glc donor, GT uses UDP-Glc. Similar to the potato glucosidase II mutant, knocking out *AtUTr1* causes increased expression of *BiP* as well as *CNX*, the gene encoding calnexin. However, despite the fact that the mutation completely abolishes *AtUTr1* expression, no morphological phenotype was found.

It has been proposed that the phenotypes of glucosidase I and II mutants are due to defects in the ER quality control mechanisms rather than deficiencies of specific N-glycosylated proteins (Vitale, 2001). Mutants affected in golgi maturation of N-glycans have been reported to have no obvious growth or development phenotype, suggesting that correct N-glycoprotein processing is not essential under normal conditions (von Schaewen et al., 1993; Strasser et al., 2007).

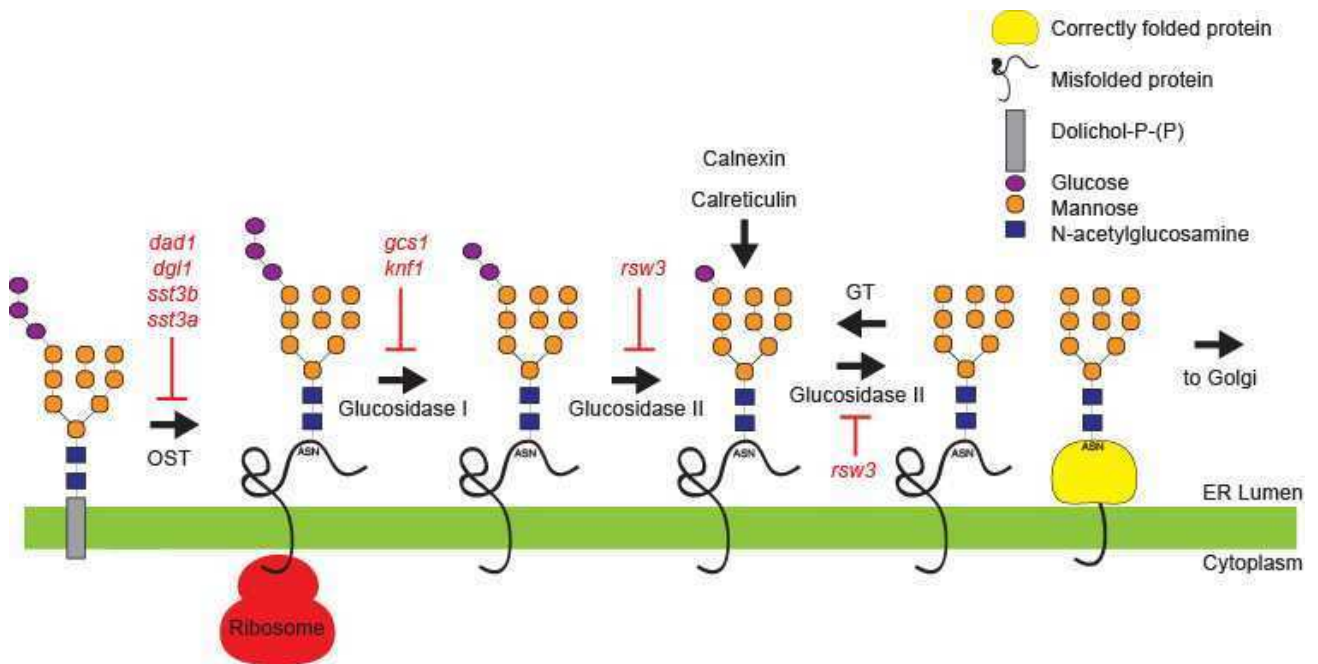


Figure 1-3. The calnexin-calreticulin cycle. Glucosidase I and II trim terminal Glc residues from newly glycosylated proteins in the ER lumen forming $\text{Man}_9\text{GlcNAc}_2$. If correct folding has been established the protein is exported to the golgi, otherwise the glycan is re-glucosylated by GT and bound by calnexin and calreticulin which retain the protein in the ER. Once the protein has been correctly folded, the glycan is no longer re-glucosylated by GT following Glc trimming and export to the golgi can occur.

1.3.2.5 N-Glycan modification

Although LLO synthesis in the ER is essentially conserved in all eukaryotes, several plant-specific N-glycan modifications occur in the golgi apparatus (Strasser et al., 2007). As a result plant N-glycoproteins may contain both $\beta(1,2)$ linked xylose and $\alpha(1,3)$ linked fucose side chains not found in other eukaryotes. These plant specific modifications can be immunogenic to mammals and significant effort has been focused on ‘humanizing’ plant-produced glycoproteins for therapeutic use.

One strategy to prevent unwanted golgi modifications has been to target therapeutic proteins for retention in the ER (Sriraman et al., 2004). However, mutation of native protein sequences may have undesired side effects. An alternative has been mutagenesis of the enzymes involved in glycan modification in the golgi. *A. thaliana* mutants have been isolated which lack specific glycosyltransferases preventing the addition of both $\beta(1,2)$ linked xylose and $\alpha(1,3)$ linked fucose (Strasser et al., 2004; Downing et al., 2006).

1.3.3 O-linked Glycosylation

O-linked glycosylation is the most heterogeneous form of protein-saccharide linkage (Wopereis et al., 2006). The O-glycan may be attached to any amino acid with a hydroxyl group (i.e. serine, threonine, tyrosine, hydroxyproline and hydroxylysine) and at least eight different core-glycans have been described ranging from single sugar residues to long repetitive polysaccharide chains (e.g. GAGs; Wopereis et al. 2006).

Many O-glycans do not contain Man, so do not require Man-P-Dol for synthesis. However, O-mannosylation, a form of O-glycosylation where Man is added to a serine or threonine residue, has been reported in yeast cell wall proteins and animal nerve, brain and skeletal muscle proteins (Strahl-Bolsinger et al., 1999)

O-mannosylation is best characterised in *S. cerevisiae* (Strahl-Bolsinger et al., 1999). Yeast cell wall mannoproteins typically have several unbranched oligosaccharide chains each consisting of four or five Man residues each. It has been demonstrated that only the first Man in O-linked glycans originates from Man-P-Dol and it is added directly

to serine and threonine residues in acceptor proteins in the ER lumen. Extension of the chain to four to five residues occurs in the golgi using GDP-Man as a sugar donor.

As far as I am aware, no examples of O-mannosylation have been described in plants. However, as mentioned above golgi-localised GDP-man transporters have been cloned (Baldwin et al., 2001; Handford et al., 2004).

O-glucosylation has been described in plants for several proteins and is a common post translational modification of plant hormones (Sembdner et al., 1994). However it is not dependent on Glc-P-Dol. Two proteins from *Zea mays*, cisZOG1 and cisZOG2, have been identified which catalyze the O-glucosylation of cis-Zeatin using UDP-Glc and not Glc-P-Dol as the Glc donor (Martin et al., 2001; Veach et al., 2003).

1.3.4 C-linked Mannosylation

C-mannosylation is the direct carbon-carbon linkage of a single Man residue to a tryptophan residue in an acceptor protein using Man-P-Dol as a Man donor (Hofsteenge et al., 1994; Doucey et al., 1998). C-mannosyltransferase activity has been successfully shown in a variety of cultured cells through the expression, purification and HPLC-analysis of mammalian RNase 2 (Krieg et al., 1997). Attempts to show C-mannosyltransferase activity in plant cells using this method were unsuccessful but this does not necessarily mean that C-mannosylation does not occur. The assay also failed for insect cells and it has been suggested that both plants and insects may have specific C-mannosyltransferase activities that fail to recognise human RNase 2 (Krieg et al., 1997).

The function of C-mannosylation remains unclear, but it is known to occur amongst several components of the complement system; part of the mammalian innate immune response which triggers cell lysis in response to pathogen attack (Furmanek and Hofsteenge, 2000).

1.3.5 GPI anchoring

Glycophosphatidylinositol (GPI) anchoring is a unique form of glycosylation used to tether proteins to the plasma membrane (Kinoshita and Inoue, 2000). GPI is attached to the C-terminal of target proteins in the ER. GPI itself consists of phosphatidylinositol (PI), a hydrophobic phospholipid inserted into membranes, and an oligosaccharide consisting of GlcN and Man that acts as a linker for attachment to target proteins. Like LLO synthesis for N-glycosylation the biosynthetic pathway for GPI anchor biosynthesis is well defined and highly conserved (Kinoshita and Inoue, 2000).

180 proteins with signal peptide sequences in *A. thaliana* are predicted to be GPI anchored (Eisenhaber et al., 2003). GPI anchored proteins (GAPs) are thought to be involved in a number of processes including cell-cell adhesion, signalling, defence and cell-wall modification (Borner et al., 2003).

1.3.5.1 GPI anchor biosynthesis

GPI anchor biosynthesis begins on the luminal side of the ER membrane with the synthesis of GlcN-acyl-PI (Kinoshita and Inoue, 2000). Firstly, a single GlcNAc residue is transferred from UDP-GlcNAc to phosphatidylinositol (PI) to form GlcNAc-PI (see figure 1-2, step 20) (Watanabe et al., 2000). GlcNAc-PI is then modified by deacetylation of the GlcNAc residue and acylation (mostly palmitoylation) of the inositol ring of PI to form GlcN-acyl-PI (figure 1-2, steps 21 and 22). GlcN-acyl-PI is then re-orientated to face the luminal side of the ER membrane by an as of yet unidentified flippase.

In the ER lumen, GlcN-acyl-PI is mannosylated using Man-P-Dol as a Man donor (Kinoshita and Inoue, 2000) (see figure 1-2, steps 24-26). After one Man residue is added it is modified by the addition of ethanolaminephosphate (EtNP). Two more Man residues are then added from Man-P-Dol. EtNP is then added to the third Man to form a mature GPI anchor (H7 form). A further EtNP residue may then be added to the second Man (H8 form). Both H7 and H8 GPI anchors can be linked to proteins through the terminal EtNP and both are found as free GPI at the cell surface.

1.3.5.2 GPI anchor biosynthesis mutants

The importance of GPI anchors in *A. thaliana* is demonstrated by the severe phenotypes of GPI anchor biosynthesis mutants. Three mutants directly affecting GPI anchor biosynthesis have been described - *pnt1*, *seth1* and *seth2*. *PNT1* encodes GPI mannosyltransferase I, the enzyme responsible for the addition of the first Man residue to the GlcNAc-PI in the ER lumen (see figure 1-2, step 24) (Gillmor et al., 2005). *SETH1* and *SETH2* encode subunits of the GPI-GlcNAc transferase complex, which catalyzes the transfer of GlcNAc from UDP-GlcNAc to PI on the cytoplasmic side of the ER membrane (see figure 1-2, step 20) (Lalanne et al., 2004).

pnt1 mutants were identified in a screen for radially swollen embryos (Gillmor et al., 2002). Phenotypes are similar to *cyt1* (see section 1.3.1.2) and include defects in cell division, delayed transition to heart stage embryos, early seedling lethality and reduced transmission through pollen (Gillmor et al., 2005). *seth1* and *seth2* mutants are both male sterile and show decreased pollen tube germination and elongation (Lalanne et al., 2004). All three mutants have increased callose deposition and *pnt1* also has decreased deposition of crystalline cellulose (Lalanne et al., 2004; Gillmor et al., 2005).

A gene from rice encoding an EtNP transferase responsible for adding EtNP to the third Man of GPI anchors (see figure 1-2, step 26), has also been cloned and shown to be induced at the transcript level by auxin and gibberellins but no mutants have been reported (Lee and Kang, 2007).

1.3.6 Importance of Man-P-Dol and Glc-P-Dol for viability

Man/Glc-P-Dol dependent glycosylation is essential for viability in plants. Several mutants in these processes have been described and their phenotypes are summarised in table 1-1. Mutants acting in both pathways often lead to embryo and/or gamete lethal phenotypes.

The phenotype of N-glycosylation mutants varies considerably according to where the mutation acts. On one hand, mutants with defects in golgi processing of N-glycoproteins show no discernible phenotype (von Schaewen et al., 1993; Strasser et al., 2007). This suggests that either the structure (or presence) of the glycan chain is not essential for the function of N-glycoproteins or that N-glycoproteins themselves are dispensable for normal growth and development. Furthermore, treatment of kaiware radish seedlings with various golgi glycoprotein processing inhibitors did not affect growth despite completely blocking the production of complex glycans (Mega, 2004).

However, when N-glycosylation is blocked in the ER, either through tunicamycin or mutagenesis, lethal phenotypes are seen. Two possible explanations have been proposed for these phenotypes. They could result from the dysfunction of a specific N-glycoprotein as has been suggested for the cellulose deficient phenotypes of *cyt1* and *knf1* (Lukowitz et al., 2001; Gillmor et al., 2002) or, as has been proposed previously, be due to indirect effects such as chaperone saturation and general ER dysfunction caused by a build up of incorrectly processed proteins in the secretory pathway (Boisson et al., 2001; Vitale, 2001).

Similarly, mutants affected in GPI anchoring result in gamete or embryo lethal phenotypes. *seth1* and *seth2* mutants demonstrate that the first glycosyltransferase reaction (see figure 1-2, step 20) in GPI anchor biosynthesis is essential for transmission through pollen (Lalanne and Twell, 2002). However, it is surprising that *seth1* and *seth2* showed almost completely abolished transmission through pollen when other GPI anchor biosynthesis mutants, *pnt1*, *cyt1* and *pmm* (T-DNA knockout), showed normal or only modestly reduced gamete transmission (Nickle and Meinke, 1998; Gillmor et al., 2005; Hoeberichts et al., 2008).

For one of these mutants, *pnt1*, this can be explained because the mutation is not completely null (Gillmor et al., 2005). However, the other mutants, *cyt1* and *pmm*, are presumably null and although embryo lethal, show no apparent reduction in transmission through pollen (Nickle and Meinke, 1998; Hoeberichts et al., 2008). Whereas *seth1* and *seth2* act directly in GPI anchor biosynthesis, *cyt1* and *pmm* affect the production of GDP-Man and therefore Man-P-Dol, the sugar donor for the first mannosylation reaction (see figure 1-2, step 24; Lukowitz et al. 2001). One possibility is that a non-mannosylated GPI anchor is sufficient for pollen. However, a GPI anchor without Man presumably cannot be attached to proteins as the GPI-protein linkage occurs via EtNP bound to the third Man residue, suggesting a possible role for free (as opposed to protein-attached) GPI. Free GPI has been identified in a number of species, and although its precise role is unclear, it has been shown that free GPI as opposed to GAPs are essential for viability in *Leishmania mexicana* (Ilgoutz et al., 1999). It may be that a modified form of free GPI (presumably GlcN-acyl-PI) is sufficient for transmission via pollen. However, this is unlikely given that Ilgoutz et al. (1999) showed that Man-P-Dol dependent mannosylation of free GPI was required in *L. Mexicana*.

Alternatively, the absence of lethal phenotypes in *cyt1* and *pmm* pollen could be explained by mother plant effects. It could be that enough GPI anchors are present in the ER of diploid pollen mother cells (*CYT1/cyt1* or *PMM/pmm*) before meiosis and carried through to haploid pollen (*cyt1* or *pmm*). Such a hypothesis was used to explain the small residual transmission of *seth1* and *seth2* (Lalanne et al., 2004). However, this does not explain the difference in phenotype between *seth1* and *seth2* and *cyt1* and *pmm*. Therefore, instead of assembled GPI anchors, *cyt1* and *pmm* pollen may inherit enough Man-P-Dol/GDP-man from pollen mother cells to maintain GPI anchor biosynthesis.

It is also surprising that the *cyt1* and *pmm* mutants did not show the same reduced male and female fertility phenotype as the N-glycosylation glucosidase I and II mutants; *rsw3*, *gcs1* and *knf1* and the OST mutants; *dgl1* and *sst3a sst3b*. Again a plausible explanation is that haploid *cyt1* and *pmm* pollen and ovules inherit enough Man-P-Dol/GDP-Man from the diploid megasporocyte and pollen mother cells.

It remains to be seen whether the other two forms of Man-P-Dol dependent glycosylation, O-mannosylation and C-mannosylation, occur in plants and what role they play in growth and development.

1.3.7. Isolation of MPDU1/LEC35

Two non-plant PQL proteins, MPDU1 and ERS1 have been implicated in glycosylation. MPDU1 was first identified in Chinese hamster ovary (CHO) cells, where it is known as LEC35, through a screen to find suppressors of mutants with altered resistance to plant derived lectins (Ware and Lehrman, 1996). Originally, the *LEC35* gene product was thought to suppress two mutations, *lec15* and *lec35*, and the cDNA clone was named *SL15* (suppressor of *lec15*; Ware and Lehrman, 1996). However, it was later found that *SL15* could only complement *lec35* and not *lec15* (Ware and Lehrman, 1998). *SL15* was later shown to originate from the same locus as the *lec35* mutation and renamed *LEC35* (Anand et al., 2001).

Table 1-1. Summary of mutants affected in Man/Glc-P-Dol dependent glycosylation in *A. thaliana*.

Mutant	Locus	Protein	Process(es)	Phenotype				Ref.
				Embryo	Pollen	Ovule	Other	
<i>pmm-12</i>	AT2G45790	phosphomanno mutase	N-Glycosylation, C-Mannosylation, O-Mannosylation, GPI Anchoring	none reported	none reported	none reported	temp. sensitive lethality	Hoeberichts et al., 2008
<i>pmm (T-DNA)</i>	AT2G45790	phosphomanno mutase	N-Glycosylation, C-Mannosylation, O-Mannosylation, GPI Anchoring	lethal	none reported	none reported		Hoeberichts et al., 2008
<i>cyt1</i>	AT2G39770	Man-1-P guanylyltransferase	N-Glycosylation, C-Mannosylation, O-Mannosylation, GPI Anchoring	lethal	none reported	none reported		Nickle and Meinke 1998
<i>vtc1</i>	AT2G39770	Man-1-P guanylyltransferase	N-Glycosylation, C-Mannosylation, O-Mannosylation, GPI Anchoring	none reported	none reported	none reported	reduced ascorbate	Conklin et al. 1999
<i>Dgl1-1</i>	AT5G66680	OST subunit	N-Glycosylation (oligosaccharide transfer)	none reported	none reported	none reported	seedling lethal	Lerouxel et al. 2005
<i>Dgl1-2</i>	AT5G66680	OST subunit	N-Glycosylation (oligosaccharide transfer)	lethal	none reported	none reported		Lerouxel et al. 2005
<i>stt3a</i>	AT5G19690	OST subunit	N-Glycosylation (oligosaccharide transfer)	none reported	none reported	none reported	salt hyper sensitive	Koiwa et al. 2003
<i>stt3b</i>	AT1G34130	OST subunit	N-Glycosylation (oligosaccharide transfer)	none reported	none reported	none reported		Koiwa et al. 2003
<i>stt3a/stt3b</i>	AT5G19690 AT1G34130	OST subunit	N-Glycosylation (oligosaccharide transfer)	N/A	abolished transmission	abolished transmission		Koiwa et al. 2003
<i>gcs1</i>	AT1G67490	glucosidase I	N-Glycosylation (Calnexin-Calreticulin cycle)	lethal	reduced transmission	Reduced transmission		Boisson et al. 2001
<i>knf1</i>	AT1G67490	glucosidase I	N-Glycosylation (Calnexin-Calreticulin cycle)	lethal	reduced transmission	normal		Gillmor et al. 2002
<i>rsw3</i>	AT5G63840	glucosidase II	N-Glycosylation (Calnexin-Calreticulin cycle)	lethal	reduced transmission	reduced transmission		Burn et al. 2002
<i>utr1</i>	AT2G02810	UDP-Glc transporter	N-Glycosylation (Calnexin-Calreticulin cycle)	none reported	none reported	none reported	Induction of ER stress markers	Reyes et al. 2006
<i>pnt1</i>	AT5G22130	GPI mannosyltransferase I	GPI anchoring; Step 24 Figure 1-2	lethal	reduced transmission	normal		Gillmor et al. 2005
<i>seth1</i>	AT2G34980	GPI-GnT subunit	GPI anchoring; Step 20 Figure 1-2	N/A	strongly reduced transmission	normal		Lalanne et al. 2004
<i>seth2</i>	AT3G45100	GPI-GnT subunit	GPI anchoring; Step 20 Figure 1-2	N/A	strongly reduced transmission	normal		Lalanne et al. 2004
<i>lec35 like (pql4/pql6?)</i>	AT5G59470 AT4G07390	Man-P-Dol Utilisation	N-Glycosylation, C-Mannosylation, O-Mannosylation, GPI Anchoring	lethal?	reduced?	reduced?	Induction of ER stress markers?	section 1.3.7.6

1.3.7.1 Isolation of *lec35* mutants and analysis of N-linked glycosylation

lec35 mutants, originally termed *pir* (processing inhibitor resistant), were isolated in a screen for mutants with changes in cell-surface oligosaccharide composition (Lehrman and Zeng, 1989). The screen utilised the plant lectin concanavalin A (ConA) which specifically binds incompletely processed cell-surface (high-mannose and hybrid type) oligosaccharides but not fully processed (complex type) oligosaccharides (Kornfeld and Kornfeld, 1985). ConA was used in conjunction with castanospermine (Cs) and swainsonine (Sw) which specifically inhibit the golgi-localised N-glycan processing enzymes glucosidase I and II (Elbein, 1987) and mannosidase II (Tulsiani et al., 1982) and prevent the conversion of high-mannose type to complex glycans.

The basis of the screen was that Cs and Sw would increase the number of high-mannose and hybrid type (incomplete) N-linked glycans at the cell surface, by blocking oligosaccharide processing, presenting more binding sites for ConA leaving the cells more susceptible to its cytotoxic effects (Lehrman and Zeng, 1989). Accordingly, it was shown that wild type cells were hypersensitive to ConA in the presence of Cs and Sw. The *lec35* mutant showed increased resistance to the combined Cs/Sw/ConA treatment suggesting that it could produce complex glycans even in the presence of the inhibitors.

Assays of glucosidase II (inhibited by Cs) and mannosidase II (inhibited by Sw) activity in the mutant cells showed that the enzymes themselves were not affected, so it was proposed that the mutant had altered glycan synthesis and by-passed the requirement of these enzymes for maturation of glycans to the complex type (Lehrman and Zeng, 1989).

HPLC analysis of radiolabelled LLOs produced by mutant cells incubated with GDP-[³H]Man showed a block in LLO synthesis at Man₅GlcNAc₂-P-P-Dol instead of progress to Glc₃Man₉NAc₂-P-P-Dol (see figure 1-2, mutants were blocked before step 11; Lehrman and Zeng 1989). However, analysis of glycopeptides from the *lec35* cells showed that Man₅GlcNAc₂ glycans were still transferred to protein.

The fact that LLO synthesis was blocked at the first Man-P-Dol dependent reaction suggested the defect could lie in the synthesis of Man-P-Dol. However, *in vitro* enzyme assays showed that Man-P-Dol synthetase activity was normal in the mutant

(Lehrman and Zeng, 1989). Furthermore, HPLC analysis following GDP-[³H]Man labelling, showed that *lec35* cells produced similar levels of Man-P-Dol as wild type *in vivo* (Zeng and Lehrman, 1990).

Another possibility was that the glycosyltransferase enzyme itself, mannosyltransferase VI, was affected by the mutation. To test this hypothesis membranes were prepared from *lec35* cells, incubated with GDP-[³H]Man, and LLOs analyzed by HPLC (Zeng and Lehrman, 1990). In these conditions *lec35* cells were not blocked at the Man₅GlcNAc₂ step of glycan synthesis but went on to produce Glc₃Man₉NAc₂ in the same way as wild type. This result showed that not only was mannosyltransferase VI not affected but that physical disruption of cells during membrane preparation was able to alleviate the *lec35* phenotype. Therefore, since both the synthesis and utilisation of Man-P-Dol were normal, it was hypothesised that the mutation affected the localisation of Man-P-Dol during synthesis of Glc₃Man₉NAc₂-P-P-Dol from Man₅GlcNAc₂-P-P-Dol, and that physical disruption during membrane preparation, allowed Man-P-Dol to be returned to its correct orientation for *in vitro* assays.

lec35 cells were also shown to be deficient in the utilisation of Glc-P-Dol (Anand et al., 2001). Under normal conditions three Glc residues are added to Man₉GlcNAc₂-P-P-Dol in the ER lumen to produce Glc₃Man₉GlcNAc₂-P-P-Dol (see figure 1-2, steps 17-19). In *lec35* mutants no glucosylated products were found. This cannot be explained by differences in the glucosylation substrate (Man₅GlcNAc₂-P-P-Dol instead of Man₉GlcNAc₂-P-P-Dol) as in *lec15* mutants, which are deficient in Man-P-Dol synthetase (and like *lec35* produce Man₅GlcNAc₂-P-P-Dol), glucosylation occurs as normal producing Glc₃Man₅GlcNAc₂-P-P-Dol. Again, the defective glucosylation phenotype of *lec35* could be reversed by performing the reactions *in vitro*.

1.3.7.2 Analysis of GPI anchoring and C-mannosylation in *lec35* mutants

Mannosylation of GPI during GPI anchor biosynthesis is also blocked in *lec35* (Camp et al., 1993). *lec35* cells accumulate GlcN-acyl-PI suggesting a block at the first mannosylation step occurring in the ER lumen (see figure 1-2, step 24). Like the N-glycan biosynthesis phenotype, disrupting cells and performing the reaction *in vitro* can

reverse the GPI anchor biosynthesis defect, again suggesting mislocalisation of Man-P-Dol.

In addition, *lec35* cells were also found to be defective in C-mannosylation, another Man-P-Dol dependent process (Anand et al., 2001). When a known target of C-mannosylation, RNase 2, was expressed in wild type and *lec35* cells, *lec35* cells showed a more than 7-fold decrease in C-mannosylation.

1.3.7.3 Possible functions of LEC35

It has been proposed that LEC35 could be the missing Man/Glc-P-Dol flippase (see section 1.3.1.1 and figure 1-2, steps 10 and 16) responsible for re-orientating Man/Glc-P-Dol so that the sugar residue faces the luminal side of the ER membrane (Anand et al., 2001). Consistent with this hypothesis, physical disruption of *lec35* cells was able to reverse the Man/Glc-P-Dol utilisation phenotype possibly because disrupting the integrity to the ER membrane allows the re-orientation of Man/Glc-P-Dol in a protein independent manner.

However, experiments with streptolysin-O (SLO) which partially permeabilises the plasma membrane but preserves intracellular membrane processes, suggests LEC35 may not have direct flippase activity (Anand et al., 2001). Although SLO permeabilised *lec35* cells could not extend Man₅GlcNAc₂-P-P-Dol to Man₉GlcNAc₂-P-P-Dol using endogenous Man-P-Dol as a donor they could when supplied with the Man-P-Dol analogue Man-P-Cit (see section 1.3.1.1). Presuming both are transferred to the ER lumen by the same flippase, this would suggest that the flippase was intact. It is possible that Man-P-Cit enters the lumen independently of a Man-P-Dol flippase but given that Man-P-Cit is not produced endogenously, it would be difficult to explain the evolution of independent transport pathways.

Other explanations for the defect include a failure to co-localise Man/Glc-P-Dol in ER subdomains associated with the production, re-orientation and utilisation of Man-P-Dol (Anand et al., 2001). The highly hydrophobic nature of the protein (seven transmembrane domains in just 247 amino acid residues) is consistent with a role involving interaction with other highly hydrophobic elements in the ER membrane. It has been suggested that LEC35 may facilitate the transfer of Man/Glc-P-Dol from the

Man/Glc-P-Dol synthetase complexes to the flippase or from the flippase to mannosyltransferase VI.

1.3.7.4 Genetic interaction between LEC35 and GPT

Interestingly, there is evidence for a genetic interaction between *LEC35* and *GPT* (Gao et al., 2007a). *GPT* catalyses the first step in N-glycan biosynthesis, the transfer of GlcNAc-1-P to Dol-P (see figure 1-2, step 1). It has been shown that, similar to *lec35* knockouts, *GPT* over-expressers accumulate Man₅GlcNAc₂-P-P-Dol instead of Glc₃Man₉GlcNAc₂-P-P-Dol (Zhu et al., 1992). It was originally thought that this could be explained by a shortage of Dol-P for Man-P-Dol and Glc-P-Dol synthesis due to excessive consumption of Dol-P-P by *GPT*, but recent reports contradict this hypothesis as total amounts of LLOs (all intermediates from GlcNAc-P-P-Dol to Glc₃Man₉GlcNAc₂-P-P-Dol) remained constant (Gao et al., 2007a).

Similarities were noticed between the phenotype of *lec35* and *GPT* over-expressers, particularly the ability to reverse the *GPT* over-expression phenotype by disrupting cells and carrying out the reaction *in vitro* (Gao et al., 2007a). Furthermore it was found that over-expression of *LEC35* in the *GPT* over-expresser background was able to avert accumulation of Man₅GlcNAc₂-P-P-Dol and restore progression to Glc₃Man₉GlcNAc₂-P-P-Dol. Therefore it appears that *GPT* over-expression interferes with *LEC35* activity. Although, the mechanism remains unclear, *GPT* may act in a dominant negative manner replacing *LEC35* in membrane sub-domains or protein complexes.

Interestingly, this offers a likely explanation for the observation by Zeng and Elbein (1995) that a tunicamycin-resistant soybean cell line isolated by gradually increasing tunicamycin concentration was found to have increased activity of both *GPT* and Man-P-Dol synthetase. Although, this observation could not be explained at the time, it could be that the increase in Man-P-Dol synthesis was a compensatory mechanism to alleviate the affect of *GPT* inhibition of *LEC35*.

1.3.7.5 LEC35 homologues in other species

LEC35 homologues have been discovered in humans and named MPDU1 (Anand et al., 2001). Mutations in human MPDU1 are responsible for congenital disorders of glycosylation type If (CDG-If) (Kranz et al., 2001; Schenk et al., 2001b). In all reported cases, mutations leading to CDG-If are not null but like *lec35* cause increased accumulation of the LLO intermediate, Man₅GlcNAc₂-P-P-Dol and reduced amounts of Glc₃Man₉GlcNAc₂-P-P-Dol, the major form of LLO found in normal cells. Human MPDU1 was shown to correct the *lec35* mutation in CHO cells (Kranz et al., 2001). The cloning of porcine MPDU1 has also been reported (Yang et al., 2005).

Surprisingly no equivalent protein has yet been described in yeast despite the fact that most of the other steps in N-glycan biosynthesis are well described. Most of the loci involved were identified through a ‘suicide screen’ by selection of mutants with increased susceptibility to radiation delivered through [³H]Man (Huffaker and Robbins, 1982). The basis of the screen was that mutants which failed to mannosylate LLO precursors would incorporate less [³H]Man and show increased tolerance to the radiation dose. Although mutations were uncovered in the majority of enzymes in the LLO biosynthetic pathway, none were found corresponding to *lec35*. However, it may be that yeast, like *A. thaliana*, contains more than one close homologue of LEC35/MPDU1 and that functional redundancy prevented the isolation of such loci in forward genetic screens.

1.3.7.6 Prediction of a *lec35*-like mutant phenotype

Given the fundamental role played by LEC35, it would be expected that plants would have a protein(s) with homologous function. Given that *AtPQL4* and *AtPQL6* are the closest *A. thaliana* homologues of LEC35 they are good candidates. If *AtPQL4* and *AtPQL6* in plants do play a similar role to LEC35/MPDU1 in mammals (i.e. allow Man/Glc-P-Dol to be used as sugar donors) it is possible to predict the phenotype of *AtPQL4/AtPQL6* mutants. If Man-P-Dol could not be utilised, mutant plants would presumably completely lack C-mannosylation and O-mannosylation and be blocked at an early stage in N-glycosylation and GPI anchoring (see figure 1-2, steps 11 and 24). As described, other mutants have been described in these processes and when null, usually

have embryo or early seedling lethal phenotypes (*cyt1*, *pnt1*, *pmm*, *knf1*, *gcs1*, *dgl1* and *rsw3*) and usually show reduced transmission through gametes (*pnt1*, *knf1*, *gcs1*, *rsw3*, *seth1*, *seth2* and *stt3a/stt3b*).

It would be expected that a *lec35*-like mutant plant would produce Man₅GlcNAc₂-P-P-Dol in the ER rather than Glc₃Man₉GlcNAc₂-P-P-Dol. The glycan would presumably still be transferred to proteins and by-pass the requirement for ER and golgi localised glucosidases and mannosidases, given the evidence from *lec35* (Lehrman and Zeng, 1989). Whilst still in the ER the Man₅GlcNAc₂ may still be glucosylated by the UDP-Glc dependent GT (see figure 1-3) but even if it is, the evidence from CHO cells is that Glc₁Man₅₋₇GlcNAc₂ acts as a poor ligand for calnexin (Ware et al., 1995). This may cause improper protein folding and ER stress and as a result lead to phenotypes similar to those described for other calnexin-calreticulin cycle mutants (*gcs1*, *knf1* and *rsw3*).

However, a *lec35*-like mutant plant would also have defects in GPI anchor biosynthesis. GPI anchoring would presumably be blocked at the first mannosylation step (see figure 1-2, step 24) preventing protein attachment. This would presumably lead to a similar phenotype as the *pnt1*, *cyt1* and *pmm-12* mutants all of which are blocked at the same step in GPI anchor biosynthesis. All three of these mutants had embryo lethal phenotypes and *pnt1* also had reduced transmission through pollen (see table 1-1).

1.3.8 ERS1

ERS1 is a PQL protein from yeast, also implicated in glycosylation, but seemingly functionally independent of LEC35/MPDU1. The closest homologue of ERS1 amongst the *A. thaliana* PQL proteins is *AtPQL5*. ERS1 was isolated as a suppressor of *erd1Δ* (Hardwick and Pelham, 1990). *erd1Δ* itself was isolated in a screen for mutants that fail to retain luminal ER proteins (Pelham et al., 1988). The screen searched for mutants that secrete a fusion protein containing an ER retention signal (invertase-GRP78-HDEL) rather than retain it in the ER lumen. Characterisation of *erd1Δ* showed that the invertase fusion also showed altered gel mobility when immuno-precipitated from *erd1Δ* and wild-type suggesting defects in glycosylation (Hardwick et al., 1990). Pulse labelling showed that whereas wild-type cells accumulated a large amount of a high molecular weight invertase corresponding to the golgi form, *erd1Δ* accumulated a low molecular

weight invertase corresponding to an incompletely processed ER form. This suggested defects during golgi extension of glycan chains. In addition the gel mobility of bands recognised with an anti-invertase antibody was increased in the *erd1* Δ mutant. This band shift was found to be dependent on N-linked glycosylation as removal of N-glycans from wild type invertase by endoH digestion resulted in an *erd1* Δ -sized band. Although the precise structure of the oligosaccharide was not determined, its susceptibility to endoH digestion suggests it is of the high-mannose type as more complex glycans are endoH resistant.

Furthermore, slight defects were found in the trafficking of a vacuolar lumen protein (Hardwick et al., 1990). Given the pleiotropic trafficking and glycosylation phenotypes, it was proposed that ERD1 is involved in maintaining appropriate conditions in golgi compartments to allow glycoprotein modification and golgi-dependent traffic to the vacuole.

ERS1 was isolated as a suppressor of *erd1* Δ during attempts to clone *ERD1* by complementation (Hardwick and Pelham, 1990). It was found that ERS1 localises to the vacuole and endosomal compartments, and that knockout mutants showed a hygromycin sensitive phenotype (Gao et al., 2005).

1.3.8.1 ers1 Δ mutant and complementation by CTNS

Interestingly, the human PQL protein, CTNS, has been reported to complement the hygromycin sensitive phenotype of *ers1* Δ (Gao et al., 2005). The precise reason for this is unknown but may be related to a common function in proton translocation (see chapter 5 section 5.5.1).

Despite the fact that both have been implicated in glycosylation, it is unlikely that LEC35/MPDU1 and ERS1 are functionally homologous. Firstly the vacuolar localisation of ERS1 is incompatible with a role in N-linked glycosylation (although ERS1 was also found in other endosomal compartments). In addition, Gao et al. (2005) reported no glycosylation defects in *ERS1* mutants. Finally, it is difficult to imagine how the over-expression of a LEC35/MPDU1-like protein could complement the *erd1* Δ phenotype which specifically affected golgi modification of glycoproteins (Hardwick et al., 1990).

1.4 STM1

STM1 is a PQL protein from the fission yeast *Schizosaccharomyces pombe* proposed to function as a G-protein coupled receptor (GPCR) involved in control of the cell cycle in response to changes in nutritional conditions (Chung et al., 2001). Although the biochemical characterisation of this protein is incomplete and there is some doubt as to whether STM1 is a genuine GPCR, genetic evidence places STM1 in the same signal transduction cascade as heterotrimeric G-proteins (Chung et al., 2001). If any of the *AtPQL* proteins are functionally homologous to STM1 it would be of considerable interest given that few components of G-protein signalling have been identified in plants. The closest *A. thaliana* relatives of STM1 are *AtPQL1*, *AtPQL2* and *AtPQL3*.

1.4.1 G-protein signalling

The classical model of G-protein signalling involves the G-protein heterotrimer consisting of α , β and γ subunits, a GPCR and a series of effector proteins (Neves et al., 2002). The $G\alpha$ and tightly associated $G\beta\gamma$ subunits are bound to the inside surface of the plasma membrane. In the resting state, $G\alpha$ and $G\beta\gamma$ are loosely associated. Upon ligand binding by a GPCR, conformational changes in the receptor promote guanine nucleotide exchange factor (GEF) activity and the exchange of GDP for GTP at the nucleotide binding domain of $G\alpha$. This in turn causes dissociation of $G\alpha$ from $G\beta\gamma$ and both may activate downstream targets. The signal is terminated by hydrolysis of GTP to GDP and re-association of $G\alpha$ with $G\beta\gamma$. GTP hydrolysis is catalysed by the intrinsic GTPase activity of $G\alpha$ but can be accelerated by regulator-of-G-protein-signalling (RGS) proteins which promote GTP hydrolysis.

1.4.2 Isolation of STM1

STM1 was isolated as a multicopy suppressor of an *S. pombe ras1* mutant and implicated in the signal transduction pathway recognising poor nutritional conditions (Chung et al., 2001). Upon starvation, *S. pombe* undergoes sexual reproduction, a process which requires a co-ordinated series of events including the recognition of nutrient conditions, cell cycle arrest at the G_1 stage of mitosis, pheromone signalling between

haploid cells of opposite mating types and morphological changes involved in the production of conjugation tubes and cell fusion (Harigaya and Yamamoto, 2007).

RAS1 plays an important signal transduction role in this process and is essential for sporulation in diploid cells and conjugation in haploids (Fukui et al., 1986). RAS1 transmits the pheromone signal perceived by a GPCR coupled to the G α , GPA1, by activating a mitogen activated protein kinase cascade and also controls cell morphogenesis through interaction with a GEF protein (Chang et al., 1994). However, RAS1-mediated pheromone signalling only occurs after cell cycle arrest at the G₁ phase. G₁ arrest is mediated by a second G α , GPA2, which acts by lowering cAMP levels upon perception of a nutrient starvation signal (Isshiki et al., 1992).

The particular *ras1* strain used to isolate *STM1* contains the wild type *RAS1* gene under the control of a thiamine repressible promoter in place of the wild type *ras1* (Chung et al., 1998). Without thiamine, *RAS1* is expressed and the cells grow normally but with thiamine, *RAS1* expression is repressed and a lethal phenotype is observed.

Over-expression of *STM1* in this mutant reversed the synthetic lethal phenotype demonstrating a genetic interaction between *STM1* and *RAS1* (Chung et al., 2001).

1.4.2.1 Functional characterisation of *STM1*

Characterisation of *stm1* knockouts and over-expressers showed defects in starvation dependent cell cycle progression. Firstly, when transferred to nutrient-poor media, *stm1* knockouts entered G₁ arrest quicker than wild type (two hours compared to four) despite growing normally on nutrient sufficient media (Chung et al., 2001). An *stm1/ras1* double mutant however, did not undergo G₁ arrest at all and continued to grow in nitrogen-depleted media. The double mutant also grew poorly in nutritionally sufficient conditions. The *ras1* single mutant did not show any difference in growth in these conditions compared to wild type.

Inducible over-expression of *STM1* showed an increase in the development of septa; the centrally placed planes defining the position at which cytokinesis occurs during mitosis (Chung et al., 2001). In addition, vegetative cell growth in nutrient rich conditions decreased. However, the promotion of septa formation was not seen when

STM1 was over-expressed in a *gpa2* (G-protein alpha subunit 2) knockout mutant background suggesting that STM1 acts upstream of GPA2 in promoting septa formation.

Given that STM1 acts upstream of GPA2 in terms of septa formation, cAMP levels in *stm1* mutants were examined (Chung et al., 2001). As mentioned above, GPA2 controls cAMP levels according to the nutritional status of the cell. *STM1* over-expressers showed a reduction in cAMP levels similar to that observed in a *gpa2* deletion mutant. However, no additive effects on cAMP levels were observed when *STM1* was over-expressed in the *gpa2* background. In addition, no affect on cAMP was observed when STM1 was over-expressed in a strain containing a mutated version of GPA2 with a permanently active GTPase domain (which had slightly higher cAMP due to permanent activation of GIT2). Together these results suggest that STM1 requires normal function of GPA2 to regulate cAMP levels. Yeast two hybrid screens showed that a physical interaction could occur between STM1 and GPA2 but not between STM1 and GPA1.

1.4.2.2 *STM1 as a GPCR?*

The physical and genetic interactions between the two proteins led to the suggestion that STM1 is the cognate GPCR for GPA2 (Chung et al., 2001). In addition *STM1* transcription was induced by nitrogen starvation. Although no ligand was identified, it was proposed that STM1 binds a nitrogen containing compound and signals through GPA2 accordingly. However, *STM1* mutants also showed phenotypes independent of GPA2. Firstly, *STM1* over-expresser mutants showed increased transcription of three meiosis specific genes, *STE11*, *MEI2* and *MAM2*. All three were also up-regulated when STM1 was over-expressed in the *gpa2* null mutant background showing that this phenotype was independent of GPA2.

There has been some controversy over whether STM1 is a genuine ligand binding receptor. As no ligand has been identified, classical biochemical assays to test if the protein is a genuine GPCR, such as testing for agonist-promoted guanine nucleotide exchange factor activity could not be carried out. In addition, despite having a typical GPCR-like topology (seven transmembrane domains with the N-terminal facing out), STM1 shares only limited sequence homology with classical GPCRs. It has been suggested that STM1 might be a pseudo-structural inhibitor that when induced by

nitrogen starvation, titrates GPA2 out of the cAMP pathway (Hoffman, 2005). Conversely, it has been argued that such a hypothesis would not explain why STM1 over-expression was unable to lower cAMP levels in cells expressing a mutant form of GPA2 with a permanently active GTPase domain (Chung et al., 2001; Shpakov, 2007). The rationale being that if it were acting in a ligand-independent way it should be able to alter cAMP levels via GPA2 in a manner dependent only on its own expression level.

1.4.3 G-protein signalling in plants

Several signalling processes in plants involve G-proteins. The plant hormones auxin, abscisic acid, gibberellins and brassinolide all show altered signal transduction in G-protein signalling mutants of *A. thaliana* and rice (Perfus-Barbeoch et al., 2004). However, *A. thaliana* has just one G α (Ma et al., 1990), one G β (Weiss et al., 1994), and two G γ proteins (Mason and Botella, 2000, 2001). In contrast humans have at least 23 G α , 5 G β and 12 G γ subunits (Temple and Jones, 2007). It is thought that plant heterotrimeric G-proteins resemble the primordial precursor and whereas metazoan G-proteins have diversified, evolutionary constraints have restricted the number of plant G-protein components. Nevertheless, plant G-proteins act in a number of distinct signal transduction processes and selective mutagenesis of individual components give rise to distinct phenotypes (Perfus-Barbeoch et al., 2004).

1.4.3.1 GPCRs in plants

It has been suggested that GPCRs may be responsible for specificity in plant G-protein signalling (Perfus-Barbeoch et al., 2004). However, no proteins in *A. thaliana* have yet been unambiguously characterised as GPCRs. In stark contrast several hundred GPCRs have been identified in animals (Temple and Jones, 2007). *In silico* methods to identify plant GPCRs are hindered by the divergent nature of the GPCR superfamily (Temple and Jones, 2007). However large scale homology studies have identified several candidates with the characteristic GPCR architecture of seven transmembrane spanning domains and an outward facing N-terminus. For example, a recent study identified 54 candidate proteins (Moriyama et al., 2006). Interestingly *AtPQL3* was one of the proteins

identified in this study. Nevertheless only two putative GPCRs have so far been identified, GCR1 (Josefsson and Rask, 1997) and GCR2 (Liu et al., 2007), and although there is some evidence to support their function as GPCRs, full biochemical proof is lacking in both cases.

Firstly, GCR1, was identified through sequence similarity to representative GPCRs from non-plant species (Josefsson and Rask, 1997). Consistent with GCR1 being a GPCR, GCR1 gain-of-function mutants are affected in several of the processes previously shown to be mediated by GPA1 including control of the cell cycle, hormone signalling and germination (Colucci et al., 2002; Apone et al., 2003). Furthermore, GCR1 and GPA1 have been shown to physically interact by split-ubiquitin and co-immunoprecipitation assays (Pandey and Assmann, 2004). However, *gcr1* and *gpa1* guard cells showed opposite phenotypes. Whereas *gpa1* was insensitive to ABA inhibition of stomatal opening, *gcr1* was hypersensitive. Opposite phenotypes argue against the hypothesis that GCR1 is a classical GPCR acting as an ABA-receptor coupled to GPA1. Instead Pandey et al. (2004) suggest that GCR1 is a negative regulator of GPA1 possibly acting by interaction with other proteins acting as ABA receptors.

A second putative GPCR, also identified on the basis of similarity to non-plant GPCRs, was shown to physically interact with GPA1 in split-ubiquitin, bimolecular fluorescence and co-immunoprecipitation assays (Liu et al., 2007). *gcr2* knockout mutants were reported to show multiple stomatal guard cell phenotypes including insensitivity to ABA-induced closure, ABA-inhibited opening and ABA-induced inhibition of potassium inward currents. Further experiments showed that ABA was able to bind to GCR2 and could promote the dissociation of $G\alpha$ from the $G\beta\gamma$ complex.

However, there has been considerable controversy over whether GCR2 is a genuine GPCR. Detailed sequence analysis suggests that the protein is more closely related to the lanthionine synthetase family than classical GPCRs and that the protein does not have membrane spanning domains (Johnston et al., 2007). There have also been suggestions that the data pointing to an interaction between GCR2 and GPA1 may be artefactual. Furthermore, no evidence was presented for agonist-promoted guanine nucleotide exchange factor activity; a classical indicator of true GPCRs.

In addition an independent study characterising GCR2 found that *gcr2* knockout lines were indistinguishable from wild type in their response to ABA (Gao et al., 2007b). Therefore, although GCR2 may be able to bind ABA and possibly GPA2, it still cannot be confirmed as a genuine GPCR.

Another GPA1 interactor, RGS1, is a regulator of G-protein coupled signalling (RGS) protein (Chen et al., 2003). Whereas GPCRs promote the exchange of GDP for GTP at the $G\alpha$ binding site, leading to heterotrimer dissociation and activation of G-protein signalling, RGS proteins facilitate the hydrolysis of GTP to GDP, re-association of the heterotrimer and negative regulation of the signalling cascade (De Vries et al., 2000). RGS1 has been shown to bind GPA1 and promote GTP hydrolysis (Chen et al., 2003). As expected for a negative regulator of G-protein signalling, *rgs1* null mutants had opposite phenotypes to *gpa1*. Whereas dark grown *gpa1* seedlings showed reduced hypocotyl elongation, *rgs1* seedlings showed increased hypocotyl length. In addition, *gpa1* seedlings are hypersensitive to sugar-induced inhibition of germination whereas *rgs1* seedlings are insensitive.

1.4.3.2 Predicted phenotype of an *stm1*-like mutant in plants

It is difficult to predict the likely phenotype of an *stm1*-like mutant in plants given the considerable diversification of heterotrimeric G-protein signalling components between plants and other eukaryotes (Temple and Jones, 2007). In addition no ligand has been identified for STM1 (Chung et al., 2001).

However, if any of the *AtPQL* proteins are functionally homologous to STM1 in G-protein signalling, it would be expected that they localise to the plasma membrane and couple GPA1, the sole $G\alpha$ subunit in *A. thaliana*. Loss-of-function mutants would presumably share at least a subset of the phenotypes seen for *gpa1* knockouts. For example, *gpa1* knockouts have been reported to show several phenotypes including rounded leaf shape, increased cell size, increased water loss and altered sensitivity to abscisic acid and auxin (Perfus-Barbeoch et al., 2004).

In addition, in *S. pombe* *STM1* functions independently of $G\alpha$ to promote the transcription of meiosis specific genes (Chung et al., 2001). It may be that a functional homologue of STM1 in plants could also act partly independently of $G\alpha$. Indeed, GCR1,

has been shown to act independently of $G\alpha$ in response to brassinosteroids and gibberellins during seed germination (Chen et al., 2004).

1.5 Objectives and thesis outline

The aim of this project is to provide a comprehensive characterisation of the *AtPQL* family, profile their sub-cellular and tissue specific localisation and assess the level of sequence and functional homology between the *AtPQL* family and the non-plant PQL proteins discussed here.

This thesis contains three results chapters. Chapter three describes sequence analysis approaches used to characterise the family *in silico* at the DNA and protein level. Relatively little has been published on the sequence characteristics of the PQL family in any species, so this chapter aims to provide a clear definition of PQL proteins, identify family members in *A. thaliana*, compare their sequences and pinpoint potentially important sequence information.

Chapter four details a variety of approaches, including real-time reverse-transcription PCR and the analysis of transgenic plants expressing promoter::GUS and GFP-fusion constructs, to demonstrate differences between family members in sub-cellular and tissue-specific localisation. I also discuss the relevance of the localisation and expression profile with respect to potential function.

Chapter five describes the functional characterisation of family. This includes heterologous expression of *AtPQL* family members in yeast, and the isolation and analysis of mutant lines for each gene. Mutants were examined at a variety of developmental stages and under a number of stress conditions to look for changes in physiology and gene expression.

Finally, chapter six contains a general discussion of all the data presented in this thesis and draws conclusions as to the likely functions and biological role of the *AtPQL* family. In addition, open questions are posed and suggestions made for further work.

Chapter 2: Materials and Methods

This chapter is split into five parts. Section 2.1 gives details of materials. Section 2.2 describes general methods for plant tissue culture and molecular biology. Sections 2.3-2.5 describe specific materials and methods for each of the three results chapters.

2.1 Materials

Except where otherwise stated all chemicals were from Sigma (Poole, UK), Fisher-Scientific (Southampton, UK) or VWR international (Poole, UK). Enzymes for cloning, cDNA synthesis and PCR were from Invitrogen (Paisley, UK). Machinery and SYBR green kits for quantitative PCR were from Stratagene (La Jolla, USA). Restriction enzymes were from Promega (Southampton, UK) or New England Biolabs (Hitchin, UK). Kits for gel extraction and plasmid miniprep, and tissue lyser equipment was from Qiagen (Crawley, UK). Custom primers were purchased from Invitrogen. DNA sequencing was carried out by the University of Dundee sequencing service (Dundee, UK). All yeast media components were purchased from Bio101 (La Jolla, USA).

2.2 General methods

2.2.1 Plant Material and Growth Conditions

Soil: To grow plants on soil, approximately 50 seeds were placed on moist compost (Levington F2; Fisons, Ipswich, UK) in a standard plant pot (65 mm x 65 mm) and placed in darkness at 4 °C for two days to ensure stratification. The pot was then placed in a controlled growth room in long-day conditions under a propagator (16 hour photoperiod, light intensity approx. 200 $\mu\text{mol m}^{-2}\text{s}^{-1}$, 22 °C/18 °C day/night temperature and 60 %/70 % day/night relative humidity). After one week seedlings were transplanted into individual pots (65 mm x 65 mm) and the propagator removed. Plants were watered once every 3-4 days until senescence.

Agar plates: For growth on agar plates, seeds were first surface sterilised. Approximately 300 seeds were placed in a centrifuge tube with 1 mL 100% ethanol for one minute and mixed by inverting the tube regularly. The ethanol was then replaced with 1 mL mild bleach solution (2.5% sodium hypochlorite and 0.01% Tween20). The seeds were mixed for 5-10 minutes by repeated inversions. The seeds were then rinsed by removing the bleach and adding 1 mL of sterile distilled water. The rinsing step was repeated five times to ensure complete removal of the bleach solution. The tube was then placed in darkness at 4 °C for two days to ensure stratification.

Growth mediums varied between experiments as indicated. For most physiological experiments, full minimal medium (referred to as “control” medium) based on that described by Arteca and Arteca (2000) was used (see table 2-1). To prepare the medium, 1000x concentrated stocks of each reagent were mixed and diluted to 1x concentration with distilled water. Sucrose was added to a concentration of 3% and pH was adjusted to pH 5.6-5.7 with 0.1 M NaOH. For solid media, 1% agar was added before autoclaving at 121 °C for 25 minutes. The media was dispensed into square (120 x 120 mm, 70 mL per plate) or circular (75 mm diameter, 25 mL per plate) tissue culture plates (Greiner Bio-One, Frickenhausen, Germany) using titration pipettes. For low sulfur medium, MgSO_4 and ZnSO_4 was replaced with MgCl_2 and ZnCl_2 . As a result the final concentration of SO_4^- was reduced from 502.34 μM to 1.96 μM (see table 2-1).

0.5x Murashige and Skoog (MS) medium was used for antibiotic selection. 2.2 g MS powder per litre was dissolved in distilled water, supplemented with 3 % sucrose and adjusted to pH 5.6-5.7 with 0.1 M NaOH. Media was solidified using 1 % agar and sterilised as described. Antibiotic (50 $\mu\text{g/mL}$ hygromycin for selection of transgenics) was added once the medium had cooled to approximately 60 °C.

Seeds were sown using a pipette to transfer seeds one-by-one onto the surface of the solidified medium. Generally, when square plates were used, 15 seeds were sown per plate and spaced evenly in a horizontal line approximately 20 mm from the top of the plate. Plates were sealed with Parafilm and placed in a controlled growth room (16 h photoperiod, light intensity approx 180 $\mu\text{mol m}^{-2}\text{s}^{-1}$, 21 °C except where otherwise stated).

Table 2-1. Reagent concentration in control and low sulfur medium.

Full Minimal "Control" Medium		Low Sulfur "-S" Medium	
Reagent	Concentration	Reagent	Concentration
KNO ₃	1.25 mM	KNO ₃	1.25 mM
Ca(NO ₃) ₂	0.5 mM	Ca(NO ₃) ₂	0.5 mM
MgSO ₄	0.5 mM	MgCl ₂	0.5 mM
FeNaEDTA	42.5 µM	FeNaEDTA	42.5 µM
KH ₂ PO ₄	0.625 mM	KH ₂ PO ₄	0.625 mM
NaCl	2 mM	NaCl	2 mM
CuSO ₄	0.16 µM	CuSO ₄	0.16 µM
ZnSO ₄	0.38 µM	ZnCl ₂	0.38 µM
MnSO ₄	1.8 µM	MnSO ₄	1.8 µM
H ₃ BO ₃	45 µM	H ₃ BO ₃	45 µM
(NH ₄) ₆ Mo ₇ O ₂₄	0.015 µM	(NH ₄) ₆ Mo ₇ O ₂₄	0.015 µM
CoCl ₂	0.01 µM	CoCl ₂	0.01 µM

2.2.2 RNA extraction and cDNA synthesis

RNA extraction: RNA was extracted from *A. thaliana* using TRIzol solution (Invitrogen) with a modified version of the manufacturers protocol. Frozen tissue was ground using a Mixer Mill MM 300 (Qiagen, West Sussex, UK) for one minute at 25 Hz. 1.6 mL of TRIzol solution was then added to the frozen powder in a 2 mL centrifuge tube and the mixture was vortexed thoroughly. Following centrifugation at 12000 x G for 10 minutes at 4°C, 1.2 mL of supernatant was transferred to a new 1.5 mL centrifuge tube containing 240 µL of chloroform. The tubes were mixed by inversion and then centrifuged at 12000 x G for 10 minutes at 4°C producing two phases. 650 µL of the upper phase was transferred to a new centrifuge tube, again containing 240 µL of chloroform. Following another centrifugation at 12000 x G for 10 minutes at 4°C, 550 µL of the upper phase was added to a new tube containing 550 µL of isopropanol. The solution was mixed by repeated inversions to precipitate RNA. The tubes were then centrifuged at 12000 x G for 10 minutes at 4°C to produce a pellet. The supernatant was discarded and the pellet vortexed in 1 mL 70% ethanol. The tubes were then centrifuged at 7500x G for 5 minutes at 4°C and the supernatant discarded. The pellet was allowed to

air dry for 10 minutes to ensure all ethanol was removed before being re-suspended in 30 μ L RNase-free water.

RNA quantity and purity was determined spectrophotometrically. 1 to 50 dilutions of RNA were prepared and scanned between 230 and 320 nm using a Lambda35 UV spectrophotometer (Perkin Elmer, Beaconsfield, UK). Absorbance readings were corrected by comparing to water. The corrected absorbance (in optical density units) at 260 nm was multiplied by 2000 to give the concentration in ng/ μ L.

cDNA synthesis: cDNA was synthesised using Superscript II (Invitrogen) according to the manufacturers instructions. A reaction mix consisting of 1 μ g RNA, 1 μ L dNTP mix (10 mM each), 1 μ L oligo dT₍₁₅₎ primer (20 μ M) in a final volume of 13 μ L, was incubated at 65 °C for 5 minutes. The reaction contents were then immediately chilled on ice and mixed with 4 μ L of the supplied 5x reaction buffer and 2 μ L of 0.1 M DTT. The reaction was then incubated at 42 °C for 2 minutes before the addition of 1 μ L (200 units) of Superscript II reverse transcriptase and further incubation at 42 °C for 50 minutes. The reaction was terminated by incubation at 70 °C for 15 minutes.

When the cDNA was used for real-time reverse-transcription PCR, RNA was first treated with RQ1 RNase-free DNase (Promega) according to the manufacturers instructions. 1 μ g of RNA was diluted to 8 μ L, mixed with 1 μ L DNase (1 unit/ μ L) and 1 μ L 10x reaction buffer and incubated at 37°C for 30 minutes. The reaction was terminated and the DNase inactivated by the addition of 1 μ L of the supplied stop solution and incubation at 65°C for 10 minutes.

2.2.3 Polymerase chain reaction (cloning, zygosity testing and expression analysis)

Standard PCR: PCR was performed using a Gene Amp 9700 thermal cycler (Perkin Elmer). Standard PCR reaction mixes (final volume 25 μ L) were set up as described in table 2-2a. Template was usually genomic DNA or cDNA. Concentration was dependent on the application and is stated in the specific materials and methods section for each chapter. Unless otherwise stated cycling conditions were as shown in table 2-2b. Primer annealing temperature and cycle number were adjusted as appropriate

for each reaction. If the expected amplicon was longer than 1 kb, extension time was increased by 1 minute per kb.

PCR products were resolved by agarose gel electrophoresis. 5 μ L of 6x loading buffer (0.25 % bromophenol blue, 0.25 % xylene xyanol, 40% sucrose) was added to 25 μ L PCR product after thermal cycling and loaded on a 1 % agarose, 0.5x TAE gel. Bands were separated at 100 mV for 25-30 minutes. After electrophoresis, gels were stained in a 0.5x TAE solution containing 0.1 μ g/mL ethidium bromide for 10 minutes. Bands were visualised under UV light using a Gel Doc 2000 scanner (Bio-Rad, Hercules, USA).

Where necessary bands were gel extracted using a QIAquick gel extraction kit (Qiagen) according to the manufacturers instructions. DNA was eluted from QIAquick columns using 50 μ L of the supplied elution buffer. DNA concentration was determined spectrophotometrically as described for RNA except that the whole eluate was loaded into the cuvette undiluted. The corrected absorbance (in optical density units) at 260 nm was multiplied by 50 to give DNA concentration in ng/ μ L.

Table 2-2. Standard reaction mix (A) and cycling conditions (B) for PCR. Cycling conditions were adjusted as stated in the text.

A

Component	Concentration	Volume (μ L)
10x ThermoPol buffer	1x	2.5
dNTP mix (10 mM each of dATP, dCTP, dGTP and dTTP)	0.2 mM each	0.5
Forward primer (20 μ M)	0.1 μ M	0.5
Reverse primer (20 μ M)	0.1 μ M	0.5
ThermoPol <i>Taq</i> DNA polymerase (5 units/ μ L)	0.02 units/ μ L	0.1
Distilled water (nucleotide free)	-	19.9
Template	variable	1

B

Cycles	Step	Temperature ($^{\circ}$ C)	Time (minutes)
1	Denaturation	94	3
30	Denaturation	94	0.5
	Primer Annealing	57	0.5
	Extension	72	1
1	Extension	72	7

Colony PCR: Following transformation (see section 2.2.6), *Agrobacterium tumefaciens* colonies were checked by colony PCR. PCR reactions were set up as described and as template, a pipette tip was used to transfer a small amount of colony directly to the PCR tube. As a negative control, a small amount of agar from an area of the plate without colonies was used.

Proofreading PCR: When PCR was used to amplify fragments for cloning, Platinum *Pfx* proofreading DNA polymerase (Invitrogen) was used instead of standard *Taq* polymerase. Extension temperature was reduced from 72 °C to 68 °C (the optimum for *Pfx*) and 10x ThermoPol buffer was replaced with the supplied 10x *Pfx* buffer.

Quantitative PCR: Real-time reverse-transcription PCR was performed using a MX4000 qPCR system with SYBR green dye (Stratagene) for real-time detection of PCR products. Target concentration was quantified by comparing amplification in cDNA samples to standards of known concentration. The composition of each qPCR reaction mix (final volume 12.5 µL), and cycling conditions, are given in table 2-3. Template was either cDNA, DNA standards or nucleotide-free water. Standards consisted of gel-purified PCR products identical to the intended amplicon. They were produced by performing standard PCR on a mixed cDNA sample using the same primers as for qPCR. PCR products were gel purified and DNA concentration determined spectrophotometrically. Purified PCR products were first adjusted to 1 pg/µL and then further diluted by ten-fold serial dilutions to produce six different standards (ranging from 1 pg/µL to 1×10^{-5} pg/µL). cDNA was diluted ten-fold before use in qPCR. Standards for each gene were sequenced to confirm amplicon identity.

SYBR green fluorescence was measured every PCR cycle at the end of the annealing step. As SYBR green fluorescence is greatly increased in the presence of double stranded DNA, product accumulation could be monitored in real time. For each reaction, the C_t (defined as the number of cycles required to reach threshold fluorescence or the point at which fluorescence can be accurately related to initial template quantity) was calculated automatically using Stratagene MX software. A standard curve was produced by plotting the C_t values for each standard dilution against the log of the initial template quantity. Initial template quantity in cDNA samples was determined by comparing C_t values to this standard curve. Each reaction was performed in duplicate.

As a control for variations in RNA quantification, reverse-transcription efficiency and template preparation, the template quantity for each gene was expressed relative to that of a stably expressed reference gene. The reference genes used in this study were *TUBULIN5* (At1g20010) and *YLS8* (At5g08290), which has recently been suggested as a high quality reference gene due to low variations in expression (Czechowski et al., 2005).

To confirm that there was only one amplicon per reaction, a dissociation curve was produced for each sample. Following each reaction, temperature was first reduced to 55 °C to ensure all products were double stranded. Products were then melted by gradually increasing the temperature in 1 °C steps. As the temperature increases, double stranded PCR products melt and SYBR green fluorescence decreases. If only one product is amplified a single large drop in fluorescence will occur at the amplicon melting temperature. As different products have unique melting points, the presence of more than one product would produce multiple drops in fluorescence. All primer pairs used in this study produced only one product under the stated conditions. Details of sample preparation for qPCR experiments are given in the individual methods sections for each chapter. Primer sequences are given in appendix A.

Table 2-3. Standard reaction mix (**A**) and cycling conditions (**B-C**) for real-time reverse-transcription PCR.

A

Component	Concentration	Volume (μL)
2x SYBR Green qPCR Master Mix	1x	6.25
Forward primer (20 μM)	0.1 μM	0.25
Reverse primer (20 μM)	0.1 μM	0.25
Distilled water (nucleotide free)	-	0.75
Template	Variable	5

B

Cycles	Step	Temperature (°C)	Time (minutes)
1	Denaturation	95	10
40	Denaturation	95	0.5
	Primer Annealing	60	1
	Extension	72	1

C

Cycles	Step	Temperature (°C)	Time (minutes)
1	Denaturation	95	1
41	Renaturation	55 + 1 °C per cycle	0.5

2.2.4 Molecular cloning and plasmid construction

All constructs used in this thesis for expression in plants were created using the Gateway recombination system (Invitrogen). Those created for expression in yeast were produced using traditional restriction enzyme/ligase based cloning. Table 2-4 lists all the plasmids produced during the course of this study.

Gateway entry clones: For all gateway, constructs an entry clone was first created using either BP clonase or the pENTR-D-TOPO system. These entry clones contain the open reading frame of *AtPQL1-6* flanked by gateway attL recombination sites (pENTR201 *AtPQL1-6* see table 2-4). Additional entry clones were created lacking the stop codon of each open reading frame for fusion of C-terminal tags (pENTR201 *AtPQL1-6 NS*). An entry clone was also created consisting of the promoter region of *AtPQL1* for promoter::GUS fusion constructs (pENTR-D-TOPO Pro_{*AtPQL1*}). BP clonase enzyme and the pENTR-D-TOPO vector construction kit were purchased from Invitrogen.

BP clonase: For constructs created using BP clonase, a two-step PCR protocol was used to produce PCR products containing the intended open reading frame flanked by attB recombination sites. This allowed cloning into the attP containing Gateway entry vector, pDONR201. For the first PCR step, forward and reverse primers were designed so that they contained 12 bases of the Gateway attB1 (forward) or attB2 (reverse) recombination sites at their 5' end and 18-24 bases specific for the fragment to be cloned at their 3' end (see appendix A). These primers were used to amplify the appropriate fragment from a mixed cDNA sample synthesised from RNA from various organs of mature plants and seedlings. PCR reactions were set up as shown in table 2-2 and products gel purified as described. Proofreading Pfx DNA polymerase was used with a modified protocol to aid the incorporation of the attB overhangs. For the first five cycles, annealing temperature was reduced to 45 °C and then increased to 55 °C for the remaining 25 cycles.

Second step PCR used 2µL of the purified PCR product from step one as template and attB1 and attB2 primers. These primers contain full-length attB sites. Standard PCR reactions were set up with proofreading DNA polymerase as described and gel purified

products used in BP cloning reactions. The reaction mix consisted of 65 ng of the gateway donor vector, pDONR201, 1 μ L BP clonase enzyme mix and 1 μ L 5x BP reaction buffer. The reaction was incubated overnight at 25 °C and terminated by adding 0.5 μ L proteinase K and incubating at 37 °C for 10 minutes. The reaction mix was used to transform TOP10 *E. coli*. Positive transformants were selected on LB agar plates containing 50 μ g/mL kanamycin and confirmed by miniprep, restriction digest and sequencing. During the BP reaction the attB sites are modified to become attL sites allowing further cloning into Gateway expression vectors.

The presence of a *ccdB* gene (lethal to most *E. coli* strains) between the attP sites in pDONR201 acts as a negative selection marker preventing the growth of *E. coli* transformed with unrecombined vector. Following recombination, the *ccdB* gene is replaced with the attB-flanked PCR fragment. For propagation of pDONR201 (and all gateway destination vectors), the *E. coli* strain DB3.1 (resistant to *ccdB*) was used.

pENTR-D-TOPO clones: Two entry clones (pENTR-D-TOPO *AtPQL1 NS* and pENTR-D-TOPO *PRO_{AtPQL1}*) were created using the pENTR-D-TOPO system. For construction of these clones, the open reading frame of *AtPQL1* or the 1kb region immediately 5' of the start codon of *AtPQL1* was amplified by PCR with Pfx proofreading polymerase as described using the primers indicated in appendix A. A mixed sample of cDNA was used as template for the amplification of the open reading frame whereas genomic DNA was used for the amplification of the promoter region. In both cases the bases CACC was added to the 5' end of the forward primer to allow cloning with pENTR-D-TOPO. PCR products were gel purified as described.

The cloning reaction was set up according to the manufacturers instructions and consisted of 4 μ L purified PCR product, 1 μ L (20 ng) pENTR-D-TOPO vector, and 1 μ L of the supplied salt solution. The reaction mix was incubated at room temperature for 30 minutes before being used for TOP10 *E. coli* transformation as described. Positive transformants were selected on LB agar plates containing 50 μ g/mL kanamycin and confirmed by miniprep, restriction digest and sequencing.

Table 2-4. Plasmids produced in this study by gateway and conventional cloning. The name of each plasmid and the chapter in which it is used is given. **A.** Gateway cloning. The names of each entry clone for each gene and the vector it was based on is given. The names of the expression clones created from each entry clone and the stated destination vector is also indicated. “NS” signifies that the open reading frame was cloned lacking a stop codon. **B.** Conventional cloning. The name of each clone is given. These plasmids were all used in chapter 5 for expression of *AtPQL* genes in yeast.

A

Gene	Entry Vector	Entry Clone	Destination Vector	Expression Clone	Used in Chapter:
<i>AtPQL1</i>	<i>pENTR-D-TOPO</i>	<i>pENTR-D-TOPO ProAtPQL1</i>	<i>pGWB3</i>	<i>pGWB3 ProAtPQL1::GUS</i>	4
	<i>pENTR 201</i>	<i>pENTR 201 AtPQL1</i>	<i>pGWB2</i>	<i>pGWB2 35S::AtPQL1</i>	5
	<i>pENTR-D-TOPO</i>	<i>pENTR-D-TOPO AtPQL1 NS</i>	<i>pGWB5</i>	<i>pGWB5 35S::AtPQL1-GFP</i>	4
<i>AtPQL2</i>	<i>pENTR 201</i>	<i>pENTR 201 AtPQL2</i>	<i>pGWB2</i>	<i>pGWB2 35S::AtPQL2</i>	5
	<i>pENTR 201</i>		<i>pGWB6</i>	<i>pGWB6 35S::GFP-AtPQL2</i>	4
	<i>pENTR 201</i>	<i>pENTR 201 AtPQL2 NS</i>	<i>pGWB5</i>	<i>pGWB5 35S::AtPQL2-GFP</i>	4
<i>AtPQL3</i>	<i>pENTR 201</i>	<i>pENTR 201 AtPQL3</i>	<i>pGWB2</i>	<i>pGWB2 35S::AtPQL3</i>	5
	<i>pENTR 201</i>		<i>pGWB6</i>	<i>pGWB6 35S::GFP-AtPQL3</i>	4
	<i>pENTR 201</i>	<i>pENTR 201 AtPQL3 NS</i>	<i>pGWB5</i>	<i>pGWB5 35S::AtPQL3-GFP</i>	4
<i>AtPQL4</i>	<i>pENTR 201</i>	<i>pENTR 201 AtPQL4</i>	<i>pGWB2</i>	<i>pGWB2 35S::AtPQL4</i>	5
	<i>pENTR 201</i>	<i>pENTR 201 AtPQL4 NS</i>	<i>pGWB5</i>	<i>pGWB5 35S::AtPQL4-GFP</i>	4, 5
<i>AtPQL5</i>	<i>pENTR 201</i>	<i>pENTR 201 AtPQL5</i>	<i>pGWB2</i>	<i>pGWB2 35S::AtPQL5</i>	5
	<i>pENTR 201</i>	<i>pENTR 201 AtPQL5 NS</i>	<i>pGWB5</i>	<i>pGWB5 35S::AtPQL5-GFP</i>	4
<i>AtPQL6</i>	<i>pENTR 201</i>	<i>pENTR 201 AtPQL6 NS</i>	<i>pGWB5</i>	<i>pGWB5 35S::AtPQL6-GFP</i>	4
	<i>pENTR 201</i>	<i>pENTR 201 AtPQL6</i>	<i>pGWB2</i>	<i>pGWB2 35S::AtPQL6</i>	5

B

Gene	Expression Clone
<i>AtPQL1</i>	<i>pDR195 ProPMA1::AtPQL1</i>
<i>AtPQL2</i>	<i>pDR195 ProPMA1::AtPQL2</i>
<i>AtPQL4</i>	<i>pDR195 ProPMA1::AtPQL4</i>
<i>AtPQL5</i>	<i>pDR195 ProPMA1::AtPQL5</i>
<i>AtPQL6</i>	<i>pDR195 ProPMA1::AtPQL6</i>

Gateway expression clones: All expression clones in this thesis were produced by combining a gateway entry clone with an attR containing destination vector using the gateway LR clonase system. The reaction mix consisted of approximately 75 ng of the appropriate entry clone, 150 ng of the appropriate destination vector, 2 μ L LR reaction buffer and 2 μ L LR clonase. The reaction was made up to a final volume of 10 μ L by adding TE buffer pH8. The reaction was incubated overnight at 25 °C and terminated by

adding 1 μ L of the supplied proteinase K solution and incubating at 37 °C for 10 minutes. The whole reaction mix was then used to transform *E. coli* as described. Positive transformants were selected on LB agar plates containing both 50 μ g/mL kanamycin and 50 μ g/mL hygromycin and confirmed by miniprep and restriction digest.

All the destination vectors used in this study are from the pGWB (Gateway Binary) series (Nakagawa et al., 2007). The features of the T-DNA cassette of each of these vectors is shown in figure 2-1. pGWB2 was used to drive over-expression of the *AtPQL* genes in their native form. The vector consists of a cauliflower mosaic virus (CaMV) 35S promoter upstream of the gateway recombination sites. pGWB5 and pGWB6 were used to create GFP fusions. pGWB5 was used to fuse an eGFP tag to the C-terminal of the open reading frame whereas pGWB6 was used to create N-terminal fusions. Both vectors contain a 35S promoter. pGWB3 was used to fuse the promoter region of *AtPQL1* to the *GUS* reporter gene. The vector contains a copy of the *GUS* reporter gene downstream of the gateway recombination site. All four contain a hygromycin phosphotransferase (*HPTII*) gene driven by the CaMV 35S promoter as a selection marker for transgenic plants.

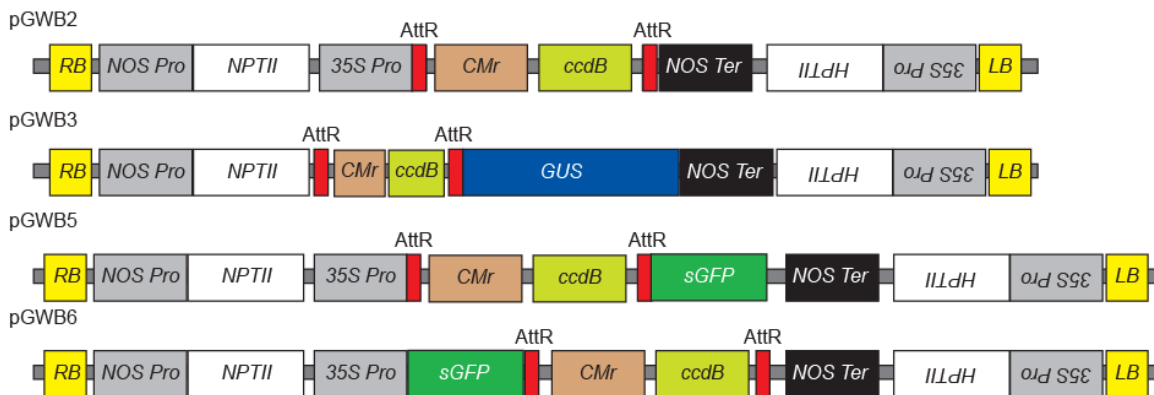


Figure 2-1. Expression cassettes of Gateway expression vectors used in this study. The T-DNA region of the pGWB2, pGWB3, pGWB5 and pGWB6 vectors are shown. In each case the Gateway attR recombination sites and plant selectable markers (*HPTII* and *NPTII* - conferring resistance to hygromycin and kanamycin respectively) are shown. The *CMr* (chloramphenicol resistance) gene is used as a selectable marker for vector propagation in *E. coli*. The *ccdB* gene (lethal to most *E. coli* strains) is used as a negative selection marker preventing the growth of *E. coli* transformed with the unrecombined vector following LR reactions. The *CMr* and *ccdB* genes are replaced with the fragment of interest (open reading frame or promoter) following recombination.

Yeast expression clones: Plasmids for the expression of *AtPQL1-6* in yeast, *pDR195 P_{PMAL}:AtPQL1-6*, were created by conventional restriction digest/ligation based cloning using the *NotI* and *BamHI* restriction sites in the multiple cloning site of the yeast episomal vector pDR195 (see figure 2-2). The open reading frame of each gene was amplified by PCR with a proofreading polymerase using approximately 500 pg of the appropriate gateway entry vector (see table 2-4) as template. Forward and reverse primers were modified to include *NotI* and *BamHI* restriction sites up and downstream of the open reading frame respectively (see appendix A).

Following gel purification of the PCR product, a restriction digest reaction was set up consisting of 50 µL purified PCR product (total eluate from column – unknown concentration), 1 µL *BamHI* (10 units/ µL), 1 µL *NotI* (10 units/µL), 6 µL reaction buffer E and 0.4 µL BSA (10 µg/µL) in a final volume of 60 µL. Reactions were incubated at 37 °C for five hours before heat inactivation at 70 °C for 15 minutes.

The episomal vector, pDR195, was prepared in the same way (4 µg was digested with *BamHI* and *NotI* as described above). Digest products were purified using QIAquick PCR purification columns (Qiagen) to remove excised fragments and the concentration quantified spectrophotometrically. Ligation reactions were set up with an approx. 3:1 ratio of insert to vector. Reaction mixes consisted of 125 ng pDR195, 50 ng purified PCR product, 4 µL 5x ligation reaction buffer and 1 µL T₄ DNA ligase (1 unit/µL) in a final volume of 20 µL. Reactions were incubated overnight at 4 °C before heat inactivation at 70 °C for 15 minutes. 5 µL of the ligation mix was used to transform *E. coli*.

Integrative plasmids for the expression of ERS1 (*pRS305 P_{ERS1}:3xMYC-ERS1*) and CTNS (*pRS305 P_{ERS1}:CTNS*) in *ers1Δ* were obtained from N. Dean (Stony Brook University).

2.2.5 *E. coli* transformation and plasmid miniprep

Transformation: *E. coli* was transformed using the thermal shock method described by Sambrook and Russell (2001). A single colony of the appropriate strain (TOP10 for cloning and DB3.1 for gateway vector preparation) was inoculated into 4 mL LB and cultured overnight to saturation at 37 °C with moderate shaking. The optical

density at 600 nm (OD_{600}) was determined by spectrophotometry and the overnight culture was diluted to an OD_{600} of 0.05 in a final volume of 250 mL LB. The culture was incubated at 37 °C with moderate shaking until it reached an OD_{600} of approximately 0.5 (2-3 hours). The culture was placed on ice for 10 minutes before being transferred to 50 mL centrifuge tubes and pelleted by centrifugation at 2500 x G for 10 minutes at 4 °C. The supernatant was replaced with 12.5 mL ice-cold 100mM $MgCl_2$. After incubation on ice for 10 minutes the cells were again pelleted by centrifugation (as above). The pellets were then resuspended in 25 mL 100mM $CaCl_2$ and left on ice for at least 30 minutes. After a final centrifugation, the supernatant was removed and the cells resuspended in 1 mL 85mM $CaCl_2$ in 15% glycerol. The cell suspension were divided into 50 μ L aliquots and frozen at -80 °C or used immediately for transformation.

To transform, cells were first defrosted on ice for approximately 20 minutes. Plasmid, ligation mix or BP/LR clonase reaction mix was then added and the cells incubated on ice for 15 minutes. They were then heat shocked at 37 °C for 45 seconds before being returned to ice for 2 minutes. 1 mL LB was then added and the cells incubated at 37 °C for one hour with moderate shaking before being plated onto LB plates (solidified with 1.5% agar) supplemented with the appropriate antibiotic (100 μ g/mL kanamycin for pDONR201/pENTR201 vectors, 100 μ g/mL kanamycin and 50 μ g/mL hygromycin for pGWB vectors, or 50 μ g/mL ampicillin for pDR195 vectors).

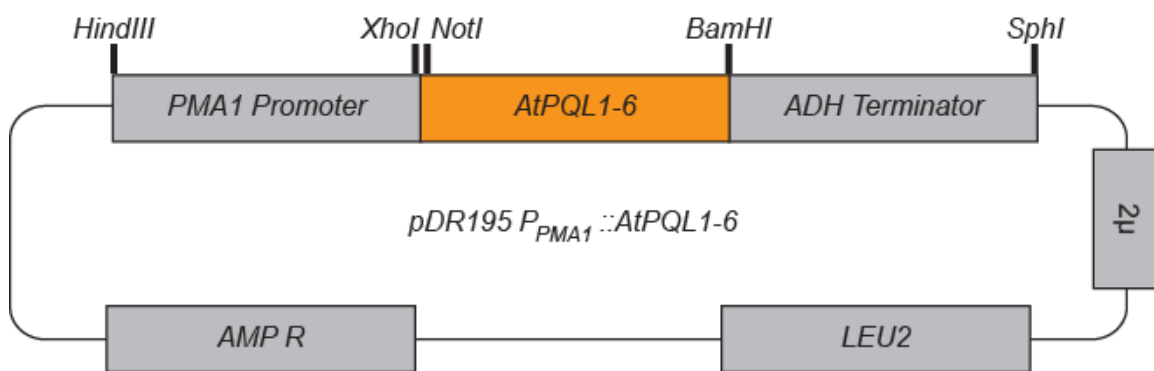


Figure 2-2. Vector map of *pDR195 P_{PMA1}::AtPQL1-6* yeast expression clones. *AtPQL1-6* was cloned between the *PMA1* promoter and *ADH1* terminator using the *NotI* and *BamHI* cloning sites in pDR195. The *SacII* site in the multiple cloning site was removed during cloning. The *AMP^R* locus (conferring ampicillin resistance in *E. coli*), *LEU2* locus (for selection on leucine-free medium) and 2 μ (yeast origin of replication) are also shown.

Plasmid Miniprep: To check potential positive colonies DNA minipreps were performed as described by Sambrook and Russell (2001). Single colonies were picked from plates with a sterile toothpick and used to inoculate 2 mL liquid culture in LB medium with the appropriate antibiotic (see above). Cultures were incubated overnight at 37 °C with moderate shaking. Once cultures became saturated they were transferred to 2 mL centrifuge tubes and pelleted by centrifugation at 15000 x G for 3 minutes. The cells were then resuspended in 100 µL resuspension buffer (50 mM sucrose, 10 mM EDTA, 25 mM, Tris pH 8) by vortexing. 100 µL lysis buffer (1 % SDS, 0.2 M NaOH) was then added to the resuspended cells and left at room temperature for 3-5 minutes before the addition of 200 µL neutralisation buffer (3 M potassium acetate pH 5). The suspension was then centrifuged for 10 minutes at 15000 x G and the supernatant transferred to new tubes containing 1 mL 100% ethanol. The tubes were then inverted repeatedly to precipitate DNA before centrifugation at 15000 x G for 10 minutes at 4 °C after which, the supernatant was replaced with 1 mL 70 % ethanol. The DNA pellet was disturbed by thorough vortexing before a final centrifugation at 15000 x G for 10 minutes at 4 °C. The pellet was allowed to air dry before resuspension in 50 µL distilled water.

To confirm the cloning reactions were successful, miniprep DNA was checked by diagnostic restriction digest. Where possible a restriction enzyme was chosen which would cut once within the cloned fragment and once in the vector backbone. Reactions consisted of 10 µL miniprep DNA, 2 µL 10x reaction buffer, 1 µL restriction enzyme (10 units/µL), 0.2 µL BSA (10 µg/µL), and 6.8 µL distilled water. Reactions were incubated at 37 °C for 2 hours and then checked by agarose gel electrophoresis.

2.2.6 *A. tumefaciens* transformation

A. tumefaciens (strain GV3101) was transformed using a thermal shock protocol. Thermocompetent cells were prepared in the same way as described for *E. coli* except incubations were performed at 28 °C instead of 37 °C and culture times were doubled (i.e. whereas *E. coli* was incubated for 2-3 hours to ensure cells were in logarithmic stage, *A. tumefaciens* was incubated for 4-5 hours). In addition, the MgCl₂ wash step was omitted,

and CaCl₂ concentration reduced from 100mM to 20mM. For culture of *A. tumefaciens*, LB was supplemented with 30 µg/mL gentamycin (selective for strain GV3101).

To transform, approximately 500 ng of vector was diluted to a total volume of 5 µL, added to the competent cells and left on ice for 15 minutes. The cells were then snap frozen in liquid nitrogen and thawed by placing in a water bath at 37 °C. They were then returned to ice for 5 minutes. 1 mL LB was then added to each tube, and the cells were incubated at 28 °C for 2 hours with moderate shaking before plating onto LB plates (solidified with 1.5% agar) supplemented with 30 µg/mL gentamycin, 100 µg/mL kanamycin and 50 µg/mL hygromycin (for all binary vectors used in this study). *A. tumefaciens* transformants were confirmed by colony PCR.

2.2.7 *A. thaliana* stable transformation

A. thaliana was transformed using the floral dip technique (Clough and Bent, 1998).

Preparation of *A. tumefaciens*: A single colony of *A. tumefaciens* containing the appropriate binary vector was inoculated into 4 mL LB and cultured overnight at 28°C with moderate shaking. This culture was then diluted into 600 mL LB and incubated at 28°C with moderate shaking for a further two days until the culture had reached saturation. The cells were then split into twelve 50 mL tubes and centrifuged at 3000 x G for 20 minutes. The supernatant was discarded and the agrobacterium pellet allowed to air dry for 10 minutes at room temperature to ensure all LB medium was removed. Each pellet was then resuspended in 4mL transformation buffer (5% sucrose, 0.5x MS medium and 0.02% Silwet L77). The suspensions were then combined and the total volume made up to 200 mL with resuspension buffer.

Preparation of *A. thaliana*: Wild-type *A. thaliana* plants were grown at a density of 12 plants per pot in standard conditions as described. After five to six weeks, all inflorescence stems were removed to promote the growth of secondary inflorescence and then used for transformation four to ten days later. Just prior to transformation, floral parts which were too old to be transformed (all open flowers and siliques) were removed.

Floral Dip: The *A. tumefaciens* suspension mix was transferred to a beaker approximately the same width as the plant pots used to prepare plants for transformation. Each pot was then placed upside-down, resting on top of the beaker with the flowers submerged in the agrobacterium suspension mix. To aid the agroinfiltration of floral parts, each pot was placed under a vacuum for five minutes. This was repeated three times per pot and afterwards the flowers rinsed briefly in distilled water to remove excess agrobacterium. The plants were returned to the growth chamber and covered with a plastic bag to maintain high humidity. After two days the plastic bag was removed and the plants left for a further two to three weeks for seed maturation.

Selection of Transformants: To select primary (T₁) transformants following floral dip transformation, seeds were harvested and allowed to dry at room temperature for at least one week in an unsealed centrifuge tube. Seeds were sown on agar plates (0.5x MS media with 50 µg/mL hygromycin) as described. Transformants could be distinguished from non-transformed plants after five to seven days as they had four green leaves and extensive root growth whereas non-transformed plants showed little or no root growth and had only two leaves which appeared chlorotic. Positive transformants were transplanted to soil for seed production.

2.3 Methods for chapter 3: alignments and sequence analysis

2.3.1 DNA and amino acid sequences

All amino acid sequences were taken from the Uniprot database (Apweiler et al., 2000). All DNA sequences, and all information regarding the positioning of introns, exons and untranslated regions were taken from the TAIR database (Rhee et al., 2003).

Diagrams of domain and intron-exon organisation were drawn to scale using Adobe Illustrator CS2 (Adobe, San Jose, USA). Positions of transmembrane domains in figure 3-4b were predicted using PHDhtm (Rost et al., 1995). The positions of PQ-loop, 7TMR-7TM, 7TMR-DISM and GGDEF domains were predicted using Pfam (Finn et al., 2006).

2.3.2 Alignment and phylogenetic tree construction

Alignments were constructed using ClustalW (Chenna et al., 2003) with a gap open penalty of 10 and a gap extension penalty of 0.2. Alignments were viewed and coloured using Jalview (Clamp et al., 2004) according to the conditions stated in the text. In figure 3-2, a consensus sequence and scores for residue conservation are given (Livingstone and Barton, 1993). Colouring by similarity is based on the Blosun62 matrix (Henikoff and Henikoff, 1992). Colouring by hydropathy is based on the Kyte-Doolittle hydropathy index (Kyte and Doolittle, 1982).

Pairwise alignments were made using the EMBOSS pairwise alignment algorithm at the EBI website (Rice et al., 2000). The needle option was chosen for alignment across the whole length of the sequences (gap open penalty 10, gap extension penalty 0.2).

The phylogenetic tree shown in was constructed using ClustalW (Chenna et al., 2003) and viewed using Treeview (Page, 1996). The alignment used for tree is shown in figure 3-5. Bootstrap values were calculated using 10,000 trials.

2.3.3 Hydropathy plots

Kyte-Doolittle hydropathy plots shown in figure 3-1 were constructed using Biology Workbench (workbench.sdsc.edu). Window size was set to eleven. Hydropathy plots in figure 3-3 are of full length proteins and were taken from the Aramemnon database (Schwacke et al., 2003). In each case transmembrane domain prediction using the PHDhtm algorithm was chosen (Rost et al., 1995).

2.3.4 Blast and Pfam searches

Blast searches were performed using BlastP (Altschul et al., 1997). In each case the full-length amino acid sequence was queried against the reference sequence database.

Pfam searches were performed using Pfam version 22.0 (Finn et al., 2006). Both global and fragment Pfam searches were chosen. In both cases, the cut-off level was set to include hits with e-values of 1 or less.

2.3.5 Gene duplication testing

To check for possible *AtPQL* gene duplication, each gene was queried against the homologous region database produced by Simillion et al. (2002) to see if they occurred on known homologous regions of the *A. thaliana* genome. To check for the presence of homologues on duplicated regions between *A. thaliana* and *O. Sativa*, *P. trichocarpa* or *S. bicolor* each locus was queried against the plant genome duplication database (Wang et al., 2006).

2.3.6 Analysis of promoter regions

For the identification of known motifs in the promoter regions of each gene, the Athena program was used (O'Connor et al., 2005). The search region set to 1 kb upstream of the transcription start site (or up to the adjacent gene if this region was shorter than 1 kb).

2.4 Methods for chapter 4: expression profile and sub-cellular localisation

2.4.1 Plant growth conditions, viral infection and wounding treatments

Wild-type (Col0) *A. thaliana* plants were used for all qPCR experiments except where otherwise stated. All harvested samples were snap frozen in liquid nitrogen and stored at -80°C.

Tissue-specific localisation: Plants were grown on soil as described. After six-weeks the specified organs were harvested. Open flowers were selected which had reached stage 13 of flower development as defined by Smyth et al. (1990). White petals were visible but anthers were still yellow (as opposed to brown) and had not yet dehisced. This should coincide with the time when mature pollen is present in anthers (Johnson-Brousseau and McCormick, 2004). Care was taken to remove the flower by cutting as close to the sepals as possible to minimise the amount of contaminating stem tissue. Approximately 10 flowers were collected from each of three different plants and pooled

for each RNA extraction. Floral buds were selected which had reached stage 6-12 of flower development (Smyth et al., 1990). Petals were still slightly shorter than sepals. Again, approximately 10 buds were collected from each of three different plants and pooled for each RNA extraction. Leaf samples were collected after removing all floral parts and harvesting the whole rosette taking care not to include any root or stem material. Rosettes from three individual plants were pooled prior to grinding in liquid nitrogen. 100-200 mg of leaf tissue was used for each RNA extraction. Roots were carefully removed from the soil and rinsed in chilled distilled water to remove as much contaminating soil as possible. Again roots from three individual plants were pooled prior to grinding in liquid nitrogen.

Diurnal time-course: To monitor diurnal changes in expression, plants were grown on agar plates as described. Twelve days after germination, shoots were harvested by carefully slicing the plant at the root-shoot apex with a scalpel and snap freezing the tissue in liquid nitrogen. Harvests were carried out at regular intervals (once every four hours) over the course of a day. For each time point, shoot material was pooled from two independent plates. Plates were placed in either long day conditions (16 h light/ 8h dark with light intensity of approximately $180 \mu\text{mol m}^{-2}\text{s}^{-1}$) or short day conditions (8 h light/ 16h dark with light intensity of approximately $220 \mu\text{mol m}^{-2}\text{s}^{-1}$).

Viral infection: RNA from plant material subjected to viral infection and wounding was provided by P. Armengaud (University of Glasgow). In both cases plants were grown hydroponically. Seeds of wild-type and *coi1-16* plants were sterilised as described and sown on sterile 1.5 mL centrifuge tubes filled full minimal medium (see table 2-1) with 1% agar. Seeds were left to germinate and grow in a sealed box in short day conditions (8 h photoperiod, approx. $200 \mu\text{mol m}^{-2}\text{s}^{-1}$, 21 °C, 70 %/60 % day/night relative humidity) for one week before being transferred to a hydroponic growth system.

The centrifuge tubes (containing one seedling each) were transferred to boxes (volume 1 L) containing liquid growth medium (see table 2-1). To allow the root to penetrate into the liquid medium the bottom 5 mm of the centrifuge tube was removed allowing contact between the liquid and agar media. The boxes were covered in black electrical tape to prevent contamination with fungi and algae. Nutrient solutions were replaced once a week with fresh media.

Plants were inoculated with cauliflower mosaic virus (CaMV) by particle bombardment, five-weeks after transfer to hydroponic medium. 5 µg DNA was diluted with water to a final volume of 50 µL, and added to a 1.5 mL centrifuge tube containing 50 µL sub-micron (0.6 µm) gold solution (Bio-Rad, Hercules, USA), 50 µL CaCl₂ (2.5 M) and 20 µL spermidine (0.1 M). The tubes were vortexed thoroughly and then left at room temperature for 2-3 minutes. They were then centrifuged briefly to pellet DNA-coated gold particles. To wash, the supernatant was replaced with 150 µL 100% ethanol and the tube vortexed thoroughly. The tubes were again briefly centrifuged and the supernatant replaced with 85 µL 100% ethanol.

Bombardment was performed using a Bio-Rad PDS-1000/He biolistic particle delivery system according to the manufacturers instructions. 1100 psi rupture discs and 5 µL of DNA-coated gold particles were used for each inoculation. As a control DNA was omitted and plants bombarded with uncoated particles. Leaf samples were collected after six days after bombardment and snap frozen in liquid nitrogen.

Wounding: Plants subjected to wounding were grown hydroponically in the same way as described for viral infection. Five-weeks after transfer to hydroponic culture, plants were wounded by piercing each leaf three times with a hypodermic needle. Plants were returned to the growth room for three hours before harvest. Wounded and non-wounded plants were placed at the opposite ends of the growth room to prevent the possible transfer of volatile signals. Whole rosettes were snap-frozen in liquid nitrogen immediately after harvest.

2.4.2 Promoter GUS fusions – transformant selection and histochemical staining

Selection of transformants: Wild-type plants were transformed with the vector *pGWB3 AtPQL1-PRO::GUS* as described. Over forty primary transformants (T₁) were initially selected on the basis of hygromycin resistance. Three of these plants, *35S::AtPQL1 PRO::GUS-1*, *35S::AtPQL1 PRO::GUS-2* and *35S::AtPQL1 PRO::GUS-3*, were selected for further characterisation and transplanted to soil and grown for seed.

The progeny (T₂) of one line *35S::AtPQL1 PRO::GUS-3* showed stunted growth and very strong GUS staining in roots. As neither affect was seen in either of the other two lines, this was thought likely to be an insertional effect. *35S::AtPQL1 PRO::GUS-3* was not included in further analysis. The T₂ progeny of the other two lines, *35S::AtPQL1 PRO::GUS-1* and *35S::AtPQL1 PRO::GUS-2* were pre-selected on hygromycin plates (to ensure each individual T₂ plant had at least one copy of the transgenic allele) for one week before being transplanted to soil and kept in long day conditions. After six weeks growth, at least nine plants of each line were harvested, separated into different organs (roots, rosette leaves, inflorescence stem, cauline leaves, flowers, isolated pollen and siliques) and stained for GUS activity.

GUS staining: GUS staining was carried out essentially as described by Jefferson et al. (1987). Tissue was harvested and immediately transferred to a petri dish containing 2 mM X-gluc solution in 50mM sodium phosphate (pH 7), and placed under a vacuum for 1 minute. The tissue was then incubated overnight in darkness at 37 °C.

The following day, tissue was rinsed several times with 50mM sodium phosphate (pH 7) and placed in increasing concentrations of ethanol to remove chlorophyll. For 20 minutes each, tissue was placed in 20%, 40%, 60%, 80% and finally 100% ethanol with gentle agitation on a rotary shaker. Stained tissue was examined with a Axiovert S100TV microscope fitted with an AxioCam HRc camera (Zeiss, Jena, Germany).

2.4.3 *AtPQL*-GFP fusion construct expression and confocal microscopy

***N. tabacum* transient transformation:** *AtPQL*-GFP fusions were transiently expressed in tobacco epidermal cells using the agroinfiltration method as described by Tang et al. (1996). *N. tabacum* (cultivar petit Havana) was grown in a growth room at 25 °C for 4-6 weeks in long-day conditions (16 hour photoperiod, light intensity approx. 200 $\mu\text{mol m}^{-2}\text{s}^{-1}$, 26 °C/22 °C day/night temperature and 60 %/70 % day/night relative humidity) until 3-4 leaves were fully expanded. A 6 mL, *A. tumefaciens* culture containing the appropriate expression construct (pGWB5 and pGWB6 vectors, see table 2-4a) was grown to saturation in LB medium by incubation for 48 hours at 28 °C with moderate shaking. The cells were pelleted by centrifugation for 5 minutes at 3000 x G

and residual LB removed by vacuum aspiration. Following resuspension in 1 mL MgCl₂ (10 mM), the OD₆₀₀ was determined and the cells diluted to an OD₆₀₀ of 0.1 in 10mM MgCl₂ and acetosyringone (1M stock in DMSO) added to a final concentration of 100 µM. The cells were then left at room temperature for 30 minutes. To infiltrate, the underside of a mature leaf was pierced with forceps and a syringe used to inject the cell suspension. The infiltrated area was marked and the plant returned to the growth room for 2-3 days before observation.

Confocal microscopy: Infiltrated areas of leaf were collected 2-3 days after infection by carefully excising an area of leaf (approximately 1 cm²) with a razor blade. To reduce background fluorescence due to air pockets, excised leaf samples were vacuum infiltrated with distilled water. Leaf discs were placed in the chamber of a 50 mL syringe filled with distilled water. With the tip blocked the plunger was drawn up to remove air bubbles from the leaf. A Zeiss CLSM510-UV microscope with a x20 Plan Apochromat objective was used to view leaf discs. GFP-fluorescence was excited at 488 nm with an argon laser. An NFT545 dichroic filter was used to split the emitted fluorescent light between two channels, with a 505 – 530 nm band-pass filter for GFP and a 560 – 615 nm band pass filter for chloroplast autofluorescence.

2.4.4 Public microarray expression database queries

Microarray database queries were performed using Genevestigator (Zimmermann et al., 2004), the Diurnal microarray viewer (Mockler, 2007), and eFP browser (Winter et al., 2007). Table 2-5 shows the Affymetrix ATH1 probe identifier for each gene checked.

Genevestigator data was collected from a variety of different microarray experiments with RNA extracted from each of the organs specified. The number of chips varies between samples. Error bars represent standard error.

The data from eFP browser was collected from microarray experiments profiling pollen at various stages of development. Pollen was collected from wild-type *A. thaliana* (ecotype *Ler* 0) plants which had been grown in a 16/8 hour light/dark cycle at 21 °C and separated into different developmental stages with a Percoll gradient. The average signal

intensity from two biological replicates for unicellular microspores, bicellular pollen and tricellular pollen is given. Data for mature pollen is from a single microarray.

The data from Diurnal was produced using RNA extracted from 7-day old seedlings grown on agar plates with 3% sucrose. Seedlings were grown at 22 °C in short or long day conditions. Short days consisted of an 8 hour photoperiod with light intensity of 180 $\mu\text{mol m}^{-2}\text{s}^{-1}$, whereas long-days consisted of a 16 hour photoperiod with light intensity of 90 $\mu\text{mol m}^{-2}\text{s}^{-1}$.

Table 2-5. Affymetrix ATH1 probe identifiers for *AtPQL*, *ALG1/2-like* and *GPT-like* genes.

Gene	Accession Number	ATH1 Probe Identifier
<i>AtPQL1</i>	<i>At4g20100</i>	254499_AT
<i>AtPQL2</i>	<i>At2g41050</i>	257357_AT
<i>AtPQL3</i>	<i>At4g36850</i>	n/a
<i>AtPQL4</i>	<i>At5g59470</i>	n/a
<i>AtPQL5</i>	<i>At5g40670</i>	249373_AT
<i>AtPQL6</i>	<i>At4g07390</i>	255199_AT
<i>ALG1/2 Like</i>	<i>At1g78800</i>	264291_AT
<i>GPT-Like</i>	<i>At3g57220</i>	251626_AT

2.5 Methods for chapter 5: functional characterisation of the *AtPQL* family

2.5.1 Yeast strains and media

The yeast strains used in this thesis are detailed in chapter 5 table 5-1. The strain *ers1Δ* (*ers1Δ::His3 : Mata ura3.52 His3 trp1-Δ901 lys2-801 suc2-Δ9 leu2-3.112*) was obtained from N. Dean (Stony Brook University, USA). All other strains were produced by transformation of *ers1Δ* with the plasmids listed in table 2-4b.

Table 2-6 gives details of the yeast media used in this study. For propagation of the *ers1Δ* strain, synthetic complete medium lacking histidine (SC-HIS) was used. *ers1Δ* contains a *His3* marker permitting growth on histidine-free media, but the strain does require uracil, tryptophan, lysine, and leucine supplements. Following transformation, positive transformants in the *ers1Δ* background were selected using SC-HIS-LEU

medium. Both the pRS305 and pDR195 plasmids contain a copy of the *Leu2* marker allowing growth on leucine free medium.

For growth tests a variety of media was used. The media was usually based on YPD with supplements added at the concentrations stated in chapter 5. To prepare yeast media, the indicated amount of each component was added to 1 L distilled water. The media was adjusted to pH 7 with 0.1 M NaOH. For solid media agar was added to a final concentration of 1%. Media was autoclaved for 20 minutes at 121 °C.

Table 2-6. Yeast media recipes. **A.** Components of synthetic complete (SC) medium. **B.** Components of YPD medium.

A

Component	Concentration (mg/L unless otherwise stated)		
	SC	SC-HIS	SC-HIS-LEU
Yeast Nitrogen Base	3.4 g/L	3.4 g/L	3.4 g/L
Ammonium Sulfate	10 g/L	10 g/L	10 g/L
Adenine	10	10	10
Arginine	50	50	50
Aspartic Acid	80	80	80
Histidine	20	0	0
Isoleucine	50	50	50
Leucine	100	100	0
Lysine	50	50	50
Methionine	20	20	20
Phenylalanine	50	50	50
Threonine	100	100	100
Tryptophan	50	50	50
Tyrosine	50	50	50
Uracil	20	20	20
Valine	140	140	140

B

Component	YPD Media Concentration (g/L)
Bacto-Yeast extract	10
Bacto-Peptone	20
Dextrose	20

2.5.2 Yeast transformation and growth tests

Transformation: Yeast were transformed using a modified version of the lithium acetate technique (Gietz et al., 1995). First a single colony was inoculated into 5 mL of synthetic complete medium supplemented with the appropriate amino acids and grown to saturation overnight at 28°C with moderate shaking. The OD_{600 nm} was checked and the cells diluted to between 0.15 and 0.2 OD units in YPD to a final volume of 50mL and grown for 3-4 hours at 28°C with moderate shaking. The cultures were then collected in sterile 50mL centrifuge tubes and pelleted by centrifugation at 1550x G at room temperature for 10 minutes. The supernatant was discarded and the cells re-suspended in 20 mL sterile distilled water. The cells were again pelleted by centrifugation (as above) and re-suspended in 20 mL TE buffer with 0.1 M lithium acetate. Following another centrifugation, the pellet was re-suspended in 200 µL TE buffer with 0.1 M lithium acetate. Competent cells were kept on ice and used within one hour for transformation.

For each transformation 7 µL of salmon sperm DNA (10 mg/mL) was denatured by heating to 95 °C for 3 minutes in a 1.5 mL centrifuge tube before being snap-cooled on ice and mixed with 1-5 µg of the appropriate plasmid, 50 µL of competent cells and 300 µL of TE buffer with 0.1 M lithium acetate and 40 % PEG4000. The transformation mix was then incubated at 28 °C with gentle shaking for 30 minutes before being transferred to a 42 °C water bath for 20 minutes. 800 µL sterile distilled water was then added to the transformation mix and an even suspension of cells obtained by inverting the tube several times. 60 µL of the transformation mix was then plated on agar plates containing the appropriate selective medium (SC-HIS-LEU).

Growth tests: Growth tests were carried out by applying serial dilutions of yeast strains to solidified media. Strains were first grown to saturation at 30 °C in liquid culture (SC media) with moderate shaking (approximately 48 hours). Each strain was then diluted to an OD₆₀₀ of 0.05 in a final volume of 50 mL YPD. The cells were then incubated at 28 °C with moderate shaking. After four hours the OD₆₀₀ was checked again and each strain diluted to 0.2 OD units. Serial dilutions were then made of each strain to give dilutions with final OD_{600s} of 0.2, 0.02, 0.002 and 0.0002. 5 µL of each dilution was applied directly to the solidified media and the plates incubated at 30 °C for two days.

2.5.3 T-DNA insertion mutant isolation

All T-DNA knockout mutants used in this study were obtained from the SALK collection via the Nottingham Arabidopsis Stock Centre (NASCC, Nottingham, UK). A mixed population of seeds (from segregating plants of the T₃ or T₄ generation) was obtained in each case. To select homozygotes, plants were grown individually on soil under and their zygosity tested by genomic DNA isolation and PCR.

Genomic DNA extraction: Genomic DNA (gDNA) was extracted using the method of Edwards et al. (1991). A leaf disc (approx 0.5 cm²) was harvested in a centrifuge tube, frozen in liquid nitrogen and ground to a fine powder using a Qiagen Mixer Mill MM 300 for one minute at 25 Hz. 400 µL gDNA extraction buffer (200 mM Tris HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5 % SDS) was added to the frozen powder in a 1.5 mL centrifuge tube. Each tube was vortexed briefly and then centrifuged at 15000 x G for 1 minute. 300 µL supernatant was removed and mixed with 300 µL isopropanol in a new centrifuge tube. The tube was vortexed thoroughly and then centrifuged for 5 minutes at 15000 x G. The supernatant was then discarded and the pellet allowed to air dry at room temperature for 15 minutes. The pellet was then resuspended in 50 µL nucleotide-free water.

Zygosity testing: Individual homozygous plants were selected using the PCR strategy suggested by SALK with primers designed using the Signal iSect tool (signal.salk.edu/tdnaprimers.2.html). The position of primers relative to the T-DNA is shown in figure 5-6. For each line, two different PCR reactions were carried out. The first, using two gene specific primers (“forward” and “reverse”, see appendix A) either side of the T-DNA insertion site was used to check for the presence of the wild-type allele. The second, using one gene-specific primer (“reverse”) and one T-DNA-specific primer (Lba1) was used to check for the presence of the mutant allele. Amplification from both indicated a heterozygous plant. Amplification from the first reaction only indicated a wild-type plant and amplification from the second reaction only indicated a plant homozygous for the T-DNA insertion. All PCR reactions were carried out using standard conditions using 1 µL gDNA as template.

Knockout confirmation: Transcript levels in two knockout lines *pql3-1* and *pql6-2* were checked by semi-quantitative reverse-transcription PCR. For *pql3-1*, RNA was extracted from a single rosette leaf of an individual plant homozygous for the T-DNA insertion and an individual wild type plant. cDNA was synthesised from both RNA samples as described and 1 µL used as a template for PCR. Standard PCR conditions were used (30 cycles) with the primers *12attB1-AtPQL3* and *12attB2-AtPQL3* (see appendix A-1). 1 µg of the vector pENTR201 *AtPQL3* was used as a positive control.

For *pql6-2*, RNA was extracted from single rosette leaves of two individual plants both homozygous for the T-DNA allele and two individual wild type plants. Reverse-transcription PCR was carried out as described for *pql3-1* with the primers *NotI-AtPQL6* and *BamHI-AtPQL6* (see appendix A). The expression of a reference gene, *ACTIN2*, was also checked using the primers *ACT2A* (CTTACAATTTCCCGCTCTGC) and *ACT2S* (GTTGGGATGAACCAGAAGGA) (Rylott et al., 2003).

To check the expression of *AtPQL6* in the line *pql6-1*, real-time reverse-transcription qPCR was used. RNA used as starting material is the same as that used for transcriptional profiling of this mutant by microarray (see section 2.5.7). The average expression values from three biological replicates are plotted. Error bars represent standard error. A two-tailed t-test was used to check for significant differences in *AtPQL6* expression between wild type and *pql6-1*.

2.5.4 Over-expresser mutant isolation

To isolate over-expressers, wild-type *A. thaliana* plants were transformed with the appropriate expression vector (pGWB2 series and pGWB5 *35S:AtPQL4-GFP* – see table 2-4) as described. The transformation rate varied from construct to construct but in each case at least five T₁ transformants were selected and transplanted to soil. The progeny of each of these plants was checked for resistance to hygromycin. Only plants showing a 3:1 ratio of resistant to sensitive offspring (indicative of a single insert) were kept for further analysis. Homozygous lines were selected from the T₂ generation. For each line, at least ten T₂ plants were grown to maturity and harvested for seeds. The zygosity of individual T₂ plants was determined by monitoring the proportion of the progeny (T₃) resistant to

hygromycin. 100 % resistant indicated a homozygous line, 75 % indicated a heterozygous parent plant and 0 % indicated a wild type parent plant (see section 5.3.4).

Semi-quantitative reverse-transcription PCR was used to check for over-expression of *AtPQL1* in the lines *35S::AtPQL1-1* and *35S::AtPQL1-2*. Semi-quantitative reverse-transcription PCR was performed as described above using the primers *AtPQL1 QL* and *AtPQL1 QR* to check the expression of *AtPQL1* and *TUB5 QL* and *TUB5 QR* to check the expression of *TUBULIN5* (see appendix A-1). As starting material RNA was extracted from a mixed tissue sample consisting of rosette leaves and open flowers. Three replicates are shown representing three individual plants of each line.

Over-expression of *AtPQL4* in the lines *35S::AtPQL4* and *35S::AtPQL4-GFP* and over-expression of *AtPQL6* in the lines *35S::AtPQL6-1* and *35S::AtPQL6-2* was checked using real-time reverse-transcription qPCR as described. RNA used as starting material for *AtPQL6* is the same as that used for the transcriptional profiling of the *35S::AtPQL6-1* and *35S::AtPQL6-2* mutants by microarray (see section 2.5.7). For *AtPQL4* over-expressers, RNA was extracted from plants treated in the same way.

2.5.5 Plant stress treatments (heat, salt, dehydration, cadmium chloride, SNP and sulfur starvation)

Heat: Sensitivity to heat shock was tested by growing plants on soil as described. For each mutant three individual seedlings were placed in the same pot as three wild-type seedlings as shown in figure 5-11. After three weeks, plants were subjected to heat shock in a 40 °C incubator overnight. To prevent dehydration, plants were well watered immediately before heat shock and then sealed in a plastic box during the treatment. Photographs were taken before heat shock, the day after heat shock, and nine days later.

Salt: Plants assayed for salt sensitivity on soil were grown in the same way as those subjected to heat shock. After three weeks growth, plants were watered with a watering solution containing 50 mM NaCl and 1/8 MS (a total of 500 mL watering solution was used for a tray containing eight 65 mm x 65 mm pots). The NaCl concentration in the watering solution was gradually increased in 50 mM intervals (see

figure 5-12) until a final NaCl concentration of 300 mM was reached 31 days after the treatment began.

A root bending assay to assess the salt sensitivity of *pql4-1* was carried out in a similar way to that described by Wu et al. (1996). In these experiments, *pql4-1* and wild-type plants were grown side by side on a plate with full minimal media under normal conditions for four days. Seedlings were then transplanted under sterile conditions to a new plate containing full minimal media or full minimal media supplemented with 100 or 200 mM NaCl. The new plates were turned upside-down so that the roots were facing up and returned to the growth room for three days before being photographed.

To check for sensitivity to salt stress applied at the point of germination, seeds were sown directly on media designed to exert salt and/or osmotic stress. Seedlings were grown exactly as normal except that media was supplemented with either 50 mM NaCl or 50 mM KCl. Plates were photographed 14 days after germination. To measure percentage water content, fresh and dry weight measurements were taken for each line. 14-day old seedlings were carefully removed from the plate with forceps and separated into roots and shoots. Roots and shoots were weighed immediately to give fresh weight (FW). They were then wrapped in foil and placed in a 70 °C drying oven for three days before being weighed again to give dry weight (DW). The percentage water content (WC%) was calculated as:

$$WC (\%) = 100 - (DW/FW \times 100)$$

Dehydration: Plants subjected to dehydration were grown in the same way as those subject to heat shock. Plants were watered every 3-4 days as normal for three weeks, after which point watering was stopped. Plants were photographed before watering was stopped, and 14 and 21 days later.

Water loss was compared from excised leaves of *pql5-1* and wild-type plants. Plants were grown on soil for four weeks in individual pots under normal conditions as described. The largest rosette leaf from three individual plants of *pql5-1* or wild type was then removed by cutting the petiole as close to the leaf base as possible using a razor blade. The leaves were placed in dry tissue culture dishes, with the lids closed but not

sealed, and left at room temperature. The leaves were weighed at regular intervals over the course of an eight hour time period.

Cadmium chloride, SNP and sulfur starvation: To check for sensitivity to CdCl₂ and SNP in control and sulfur-deplete conditions, seeds were sown directly on either control or low sulfur media (see table 2-1) supplemented with either 0, 25 or 75 μ M CdCl₂ or 0, 25 or 75 μ M SNP. Plants were grown as described except that plates with SNP were placed under low intensity light (approx 70 μ mol m⁻²s⁻¹ to prevent SNP degradation).

2.5.6 Pollen isolation, *in vitro* germination and FDA viability staining

In vitro pollen germination assays were carried out by applying pollen to pollen germination media (PGM) and incubating overnight in humid conditions. PGM consisted of 5 mM MES, 1 mM KCl, 10 mM CaCl₂, 0.8 mM MgSO₄, 1.5 mM boric acid, 16.6 % sucrose, 3.65 % sorbitol, and 10 μ g/mL myo-inositol (Fan et al., 2001). 1 % agarose was added to solidify the medium. Autoclaving was avoided to prevent caramelisation and instead agarose was dissolved by microwaving for 15-20 seconds. Media was dispensed into several 3 cm diameter tissue culture dishes (1.5 mL of media per dish).

Pollen was harvested as described by Johnson-Brousseau and McCormick (2004). Flowers were chosen which were still closed but where petals were visible coinciding with stage 13 of pollen development (Smyth et al., 1990). Using forceps, gentle pressure was applied to the base of the sepals to encourage the flower to open fully. Individual stamens were then removed using forceps at the base of the filament. Anthers were repeatedly dabbed onto solid pollen germination medium to release pollen grains. In each case only one flower was used per tissue culture dish. To ensure high humidity, a humidity chamber was constructed from an empty pipette tip box as described by Johnson-Brousseau and McCormick (2004). The tissue culture dishes were placed inside the humidity chamber and incubated overnight at 24 °C.

To assess pollen viability, fluorescein diacetate (FDA) was used essentially as described by (Heslop-Harrison and Heslop-Harrison, 1970). PGM was prepared as described above and applied to a microscope slide. Pollen was isolated directly onto the

slide and placed in a humidity chamber. To stain, 100 µL of FDA staining solution was applied directly to the slide on the under-side of a cover slip. To prepare FDA staining solution, FDA was first dissolved in acetone (2 mg/ml) and then added drop by drop to pollen germination media until the solution became turbid. The slides were incubated in a humidity chamber at 24 °C for at least 30 minutes before viewing.

FDA staining was visualised by confocal microscopy using a Zeiss CLSM510-UV microscope with a x10 Plan Apochromat objective. Standard settings for FITC were used. Fluorescence was excited at 488 nm with an argon laser set to 10% and a 505 – 550 nm band-pass filter was used for detection of FDA fluorescence.

2.5.7 Transcriptional profiling (microarray experiments)

2.5.7.1 Sample preparation, array hybridisation and scanning

Plant material for microarray analysis was produced by growing plants (wild-type, *pql4-1*, *pql4-2*, *pql6-1*, *35S::AtPQL6-1*) on agar plates under standard conditions as described. Each replicate consisted of RNA extracted from shoot tissue pooled from two individual plates each containing 14 seedlings per line. Biological replicates were sown one week apart but otherwise treated identically. All harvests were performed between 10 a.m. and 12 noon, 12 days after germination.

Microarray hybridisation and scanning was performed by the SHWFG facility at the University of Glasgow. Briefly, RNA quality was confirmed using an Agilent (Austin, USA) RNA BioAnalyzer 2100 before being reverse transcribed to double-stranded cDNA. Double stranded cDNA was then *in vitro* transcribed and biotinylated. Biotinylated cRNA was hybridised to Affymetrix (High Wycombe, UK) Arabidopsis ATH1 genes. Chips were washed and stained using the Affymetrix Fluidics 400 station before scanning with an Affymetrix Gene Array 2500 scanner.

The up-regulation of *BiP3* was confirmed using real-time reverse-transcription PCR as described. Plant material was the same as that used for microarray analysis. *BiP3* primer sequences were taken from Reyes et al. (2006).

2.5.7.2 Data analysis

Data was analyzed using the automated Funalyse (version 3.3.6) analysis software (SHWFG facility, University of Glasgow). This analysis consisted of the robust multichip average (Irizarry et al., 2003) normalisation followed by rank product (Breitling et al., 2004b) and iterative group analysis (Breitling et al., 2004a).

Rank product (RP): Rank product analysis is a method for identifying differentially expressed genes (Breitling et al., 2004b). Genes are first ranked according to fold change in each individual replicate, after which the individual ranks are then multiplied to give the RP. The entire gene list is then sorted by RP and each gene is assigned a p-value based on the likelihood of a rank product as high or low as observed being achieved at random.

Iterative group analysis (iGA): iGA is a method to detect changes in expression of groups of genes pre-assigned to functional categories (Breitling et al., 2004a). Genes are automatically grouped according to various criteria including gene ontology assignments, TAIR gene families, key words extracted from gene descriptions in literature and BLAST search results. The iGA procedure determines groups that are enriched at the top of gene lists sorted by RP (separately for up- and down-regulation). A hypergeometric test is used to assign a p-value to each group. iGA iteratively determines the subset of genes in each group which minimise the p-value.

Sub-cellular localisation: The predicted sub-cellular localisation of each differentially regulated gene was determined using TargetP (Emanuelsson et al., 2000) on the TAIR website (www.arabidopsis.org/tools/bulk/protein/index.jsp). For each protein TargetP determines four scores based on the likelihood that the protein localises to the chloroplast, mitochondria, secretory pathway or any other location (eg cytosol). The compartment with the highest score is given as the most likely localisation for each protein and is indicated in appendix B Where the difference between the highest and second highest scores is less than 0.4 a value of undefined is returned.

To determine the proportion of genes in the *A. thaliana* genome encoding proteins predicted to be targeted to the secretory pathway each accession in the genome was analyzed using the above method. Of the 31,917 genes in the TAIR database, 2723

(8.5%) are predicted to be targeted to the chloroplast, 441 (1.4%) to the mitochondria, 5540 (17.4%) to the secretory pathway and 7282 (22.8%) to another location. 15931 (49.9%) are of undefined sub-cellular localisation. These figures were used as the expected values for χ^2 tests.

N-glycosylation status: The predicted N-glycosylation status of each protein sequence was determined using NetNGlyc 1.0. The attachment site for N-glycans is known as Asn-X-Ser/Thr (where X is any amino acid except proline). However, not all Asn-X-Ser/Thr motifs are actually glycosylated and whether they are depends on a number of factors including the amino acids after the motif and at the X position (Shakin-Eshleman et al., 1996; Mellquist et al., 1998). In addition, only proteins that are trafficked through the secretory pathway can be N-glycosylated making the presence of a signal peptide an essential pre-requisite.

For each protein sequence, every Asn-X-Ser/Thr motif is assessed by using NetNGlyc 1.0 and rated as -, +, ++ or +++. Those proteins with one or more + rated N-glycosylation sites and a predicted signal peptide were considered probable N-glycoproteins and are indicated in appendix B.

To determine the approximate rate of N-glycosylation in *A. thaliana* a random sample of 300 accession numbers (chromosome number 1-5 and co-ordinates 00000-50000 in integers of 10) was generated using a random number generator (www.random.org). 229 were genuine *A. thaliana* accession numbers of which 35 (16%) encoded predicted N-glycoproteins. This figure was used as the expected value for the χ^2 test.

In addition, the predicted protein sequences of differentially regulated genes with altered expression in each mutant were checked for the presence of GPI anchor attachment sites. In total, 180 *A. thaliana* proteins (0.94 % of the genome) are predicted to contain both a signal peptide and a predicted GPI anchor attachment site, both of which are prerequisites for GPI anchor attachment (Eisenhaber et al., 2003). Of these 180, five showed differential regulation in the microarray experiments and are indicated in appendix B.

Chapter 3: Alignments and Sequence Analysis

3.1 Introduction

There are currently few descriptions of the PQ-loop repeat (PQL) family in the literature. Apart from reports of characterisation of individual proteins (see chapter one), two papers have previously described characteristics of the family. Ponting et al. (2001) first identified the PQ-loop domain whilst searching for *D. melanogaster* proteins with internal repeats. They describe the domain as having two transmembrane helices separated by a loop. Having two copies of this domain defines a protein as a member of the PQL family.

Zhai et al. (2001) also identified the family (referred to as LCT - lysosomal cystine transporter family) whilst looking for distant homologues of the microbial rhodopsins (MR) and predict the two families share common ancestry. Partial sequence alignment of members of the MR family with members of the PQL family from yeast showed considerable sequence conservation. Like Ponting et al (2001) they identify internal repeats and suggest the family arose from a gene duplication event. They identify two full-length *A. thaliana* proteins (referred to here as *AtPQL1* and *AtPQL3*) as members of this family.

In this chapter I aim to clarify the definition of the PQL family, identify and describe members from *A. thaliana* (the *AtPQLs*), and analyse the sequence of these proteins. I will discuss characteristic features of the family and make predictions regarding their functional relevance. I will also compare the *AtPQL* sequences with their non-plant homologues, especially where functional information is available. In particular, I will look at the level of conservation in specific residues and motifs, identified as being critical for function in non-plant PQL proteins. The analysis in this chapter pays particular attention to ten PQL proteins, *AtPQL1-6* of *A. thaliana*, CTNS, MPDU1 (the human homologue of LEC35), ERS1 and STM1 (see table 3-1).

Finally, I will examine the nucleotide sequences of the *AtPQL* genes. This will include analysis of their genome distribution, intron-exon organisation and predicted promoter sequences.

3.2 Results

3.2.1 The PQ-Loop domain

As mentioned previously, the defining feature of the PQL family is the presence of two copies of the PQ-loop domain. This domain is variously referred to as the PQ-loop (Pfam PF04193; Finn et al., 2006), CTNS domain (Smart SM00679; Schultz et al., 1998) and Cystinosin/ERS1p repeat (Interpro IPR006603; Apweiler et al., 2006). In this thesis, this domain will be referred to as the PQ-loop.

The PQ-loop was first identified in the *D. melanogaster* protein CG17119 in a database search for repeated regions within single proteins (Ponting et al., 2001). The domain consists of two transmembrane helices and an intervening hydrophilic core. The PQ-loop takes its name from a highly conserved proline-glutamine dipeptide motif close to the centre of the domain.

The Pfam database contains a total of 346 known or predicted proteins containing the PQ-loop of which 336 are from eukaryotic species (Finn et al., 2006). An alignment of the PQ-loops of the ten PQL proteins listed in table 3-1, demonstrates their sequence similarity (see figure 3-2a). Like *DmCG17119*, the vast majority of proteins containing a PQ-loop have two copies of the domain. Such proteins, will be referred to as PQ-loop repeat or PQL proteins.

The best conserved residues form the PQ-dipeptide motif at position 22-23. This motif is present in almost all PQ-loops. Notable exceptions include the yeast protein, ERS1 which has two proline residues instead of a proline and a glutamine in the first PQ-loop and the human protein, MPDU1 which has a leucine instead of a proline in the second PQ-loop.

The conservation however, is not restricted to the PQ-dipeptide. The PQ-loop consists of around 55-70 amino acid residues. Other particularly strongly conserved residues include non-polar hydrophobic amino acids (isoleucine, valine, leucine or phenylalanine) at positions 41 and 55.

Hydropathy analysis shows the domain consists of two hydrophobic regions surrounding a hydrophilic core positioned around residues 29-32 (see figure 3-2b). This

hydrophilic core usually has a serine at position 32 preceded by two or three highly hydrophilic residues such as lysine or arginine. Hydropathy plots of selected PQ-loops (figure 3-1) and of full-length PQL proteins (figure 3-3) show that the PQ-loop usually spans two transmembrane domains.

Table 3-1. Names and accession numbers of PQL proteins referred to in this chapter.

Organism	Name	AGI Accession	UniProt Accession	No. of AA Residues	Position of 1st PQ-loop	Position of 2nd PQ-loop
<i>H. sapiens</i>	CTNS	N/A	O60931	367	123-189	263-328
<i>H. sapiens</i>	MPU1	N/A	O75352	247	39-105	159-216
<i>S. pombe</i>	STM1	N/A	Q10482	271	14-80	185-239
<i>S. cerevisiae</i>	ERS1	N/A	P17261	260	1-67	162-212
<i>A. thaliana</i>	AtPQL1	At4g20100	O49437	288	3-69	187-242
<i>A. thaliana</i>	AtPQL2	At2g41050	Q8RXY4	376	8-74	268-332
<i>A. thaliana</i>	AtPQL3	At4g36850	O23198	374	30-96	255-321
<i>A. thaliana</i>	AtPQL4	At5g59470	Q9LTI3	239	27-93	150-205
<i>A. thaliana</i>	AtPQL5	At5g40670	P57758	270	9-75	151-213
<i>A. thaliana</i>	AtPQL6	At4g07390	Q8VY63	234	27-93	150-205

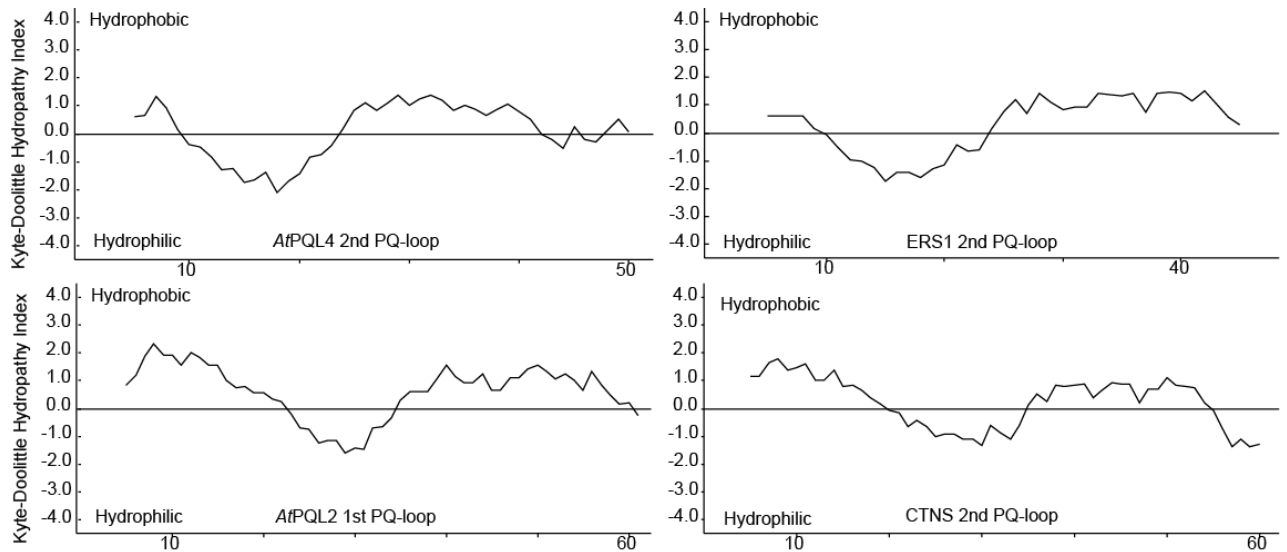


Figure 3-1. Hydropathy plots of selected PQ-loops. Hydropathy plots for the amino acid sequences of selected PQ-loops were constructed using the Kyte-Doolittle method with a window size of eleven.

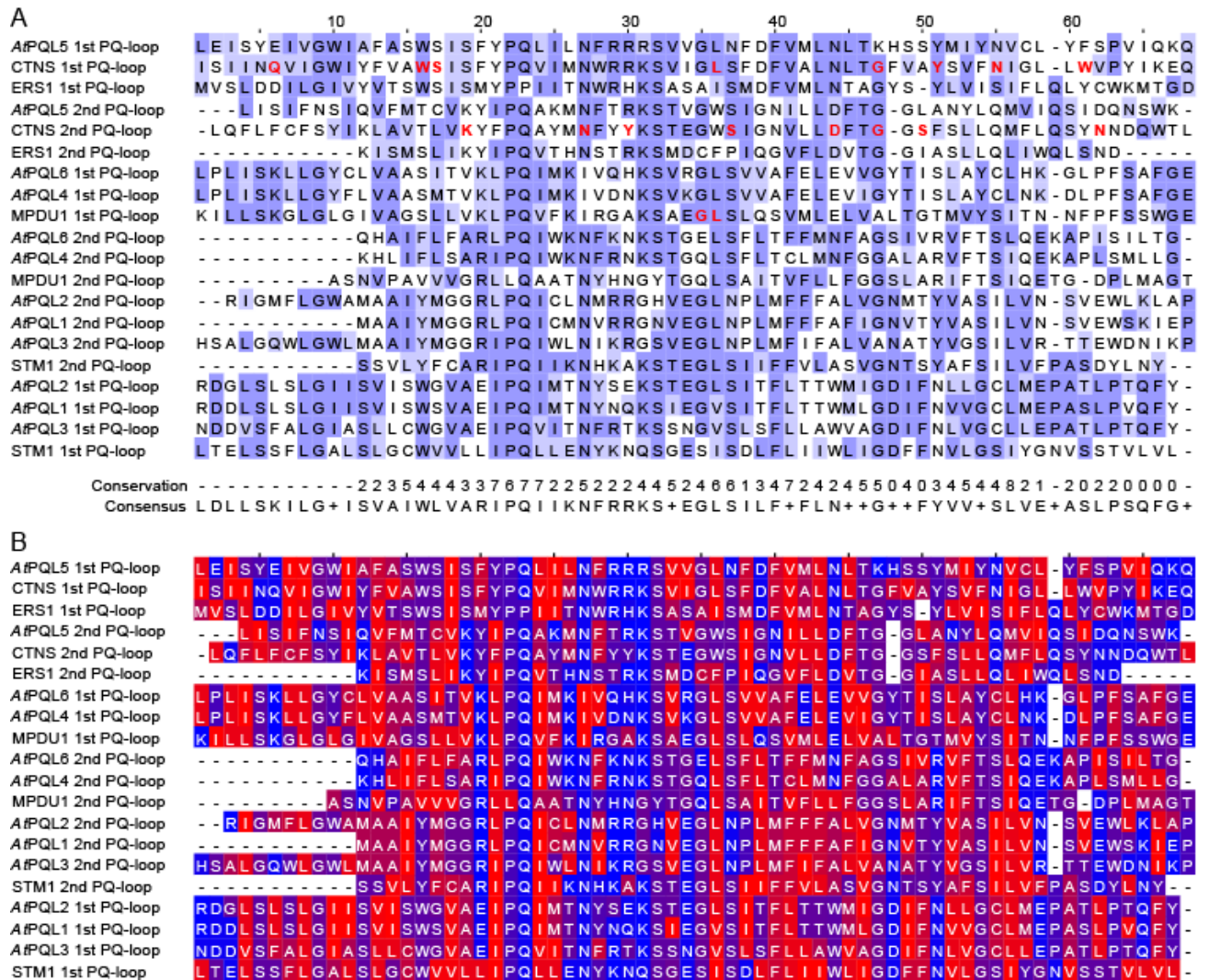


Figure 3-2. Sequence alignment of PQ-loops from AtPQL1-6, CTNS, ERS1, MPDU1 and STM1. Alignments were coloured according to similarity (**A**) or hydrophobicity (**B**). **A.** Residues matching the consensus sequence are coloured dark blue. Similar residues (i.e. positive Blosum62 score) are coloured light blue. Affected residues in known mutant forms of human CTNS and MPDU1 are highlighted in red. Residue conservation scores are displayed (Livingstone and Barton, 1993). The columns showing the greatest conservation of physico-chemical properties score highest. **B.** Residues are coloured according to the Kyte-Doolittle hydrophathy index. Colours range from red (most hydrophobic) to blue (most hydrophilic).

3.3.2 The PQ-loop repeat family in *A. thaliana*

The Pfam database shows six *A. thaliana* proteins containing two copies of the PQ-loop (Finn et al., 2006). These six proteins will be referred to as *AtPQL1-6*. No *A. thaliana* proteins were found with just one copy of the PQ-loop. Although the Pfam database does show *A. thaliana* proteins with only one PQ-loop, this is the result of partial sequence information from incomplete cDNA clones. For example, Q8GZF5 results from a partial clone of *AtPQL6* (Hall et al., 2003).

Hydropathy plots of *AtPQL1-6* are shown in figure 3-3. Each is predicted to have seven transmembrane-spanning domains with the N-terminus facing the non-cytoplasmic side of the membrane. The presence of seven predicted transmembrane segments is common amongst all the PQL proteins characterised so far (Hardwick and Pelham, 1990; Town et al., 1998; Anand et al., 2001; Chung et al., 2001).

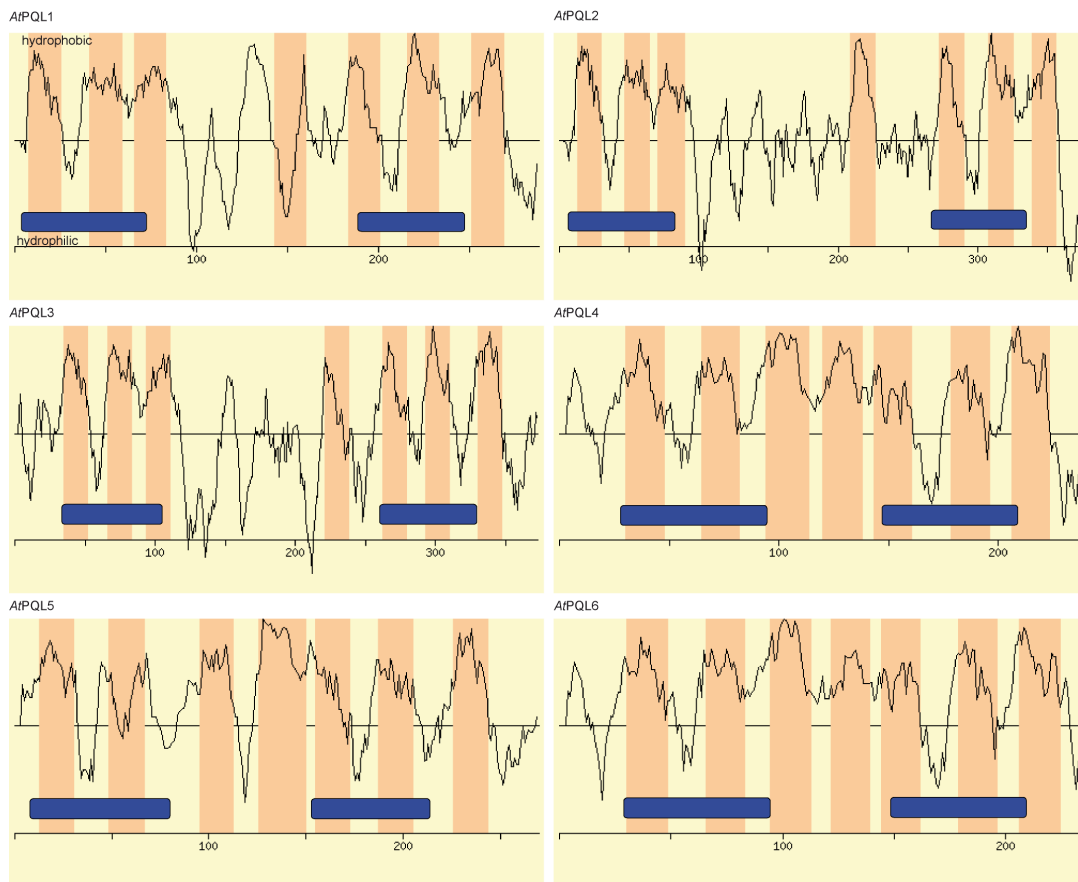


Figure 3-3. Hydropathy plots of *AtPQL1-6*. Plots are from the Aramemnon database (Schwacke et al., 2003) with transmembrane segments predicted using PHDhtm (Rost et al., 1995). Orange represents transmembrane regions and blue bars indicate PQ-loops.

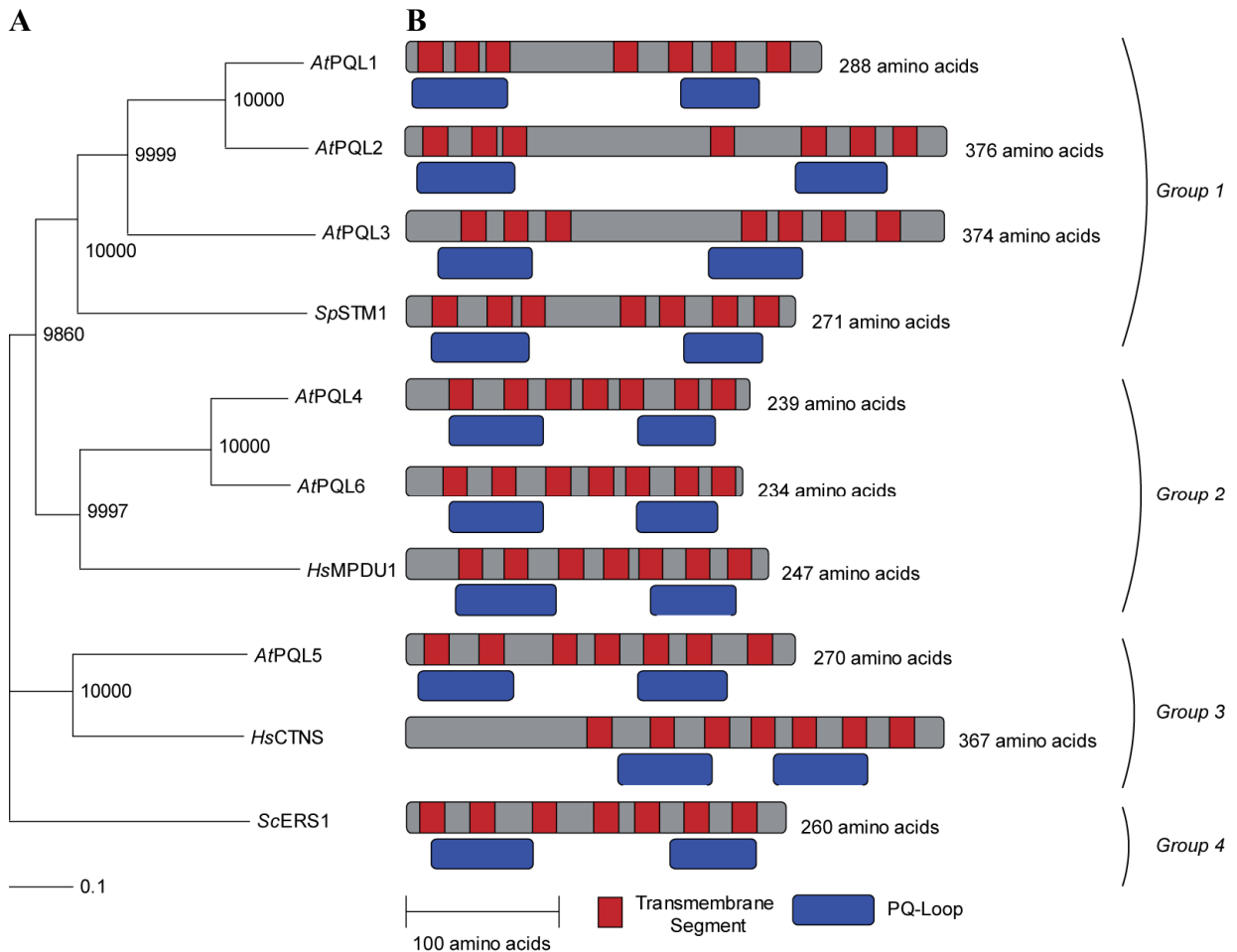


Figure 3-4. Phylogenetic tree and domain organisation of selected PQL proteins. The phylogeny (**A**) was created from a full-length alignment of *AtPQL1-6*, CTNS, MPDU1, STM1 and ERS1 (see figure 3-5). Tree was created using ClustalW with node numbers indicating bootstrap values (10,000 trials). The distribution of transmembrane segments and PQ-loops in each protein is shown (**B**). The sub-groups of family members referred to in the text are indicated here.

3.3.2.1 PQL family sub-groupings

Figure 3-4a shows a phylogenetic tree to illustrate the relationships between *AtPQL1-6* and those PQL proteins from non-plant species where functional information is available. The functions of these proteins are described in detail in the introduction. Briefly, CTNS is a human lysosomal cystine transporter (Kalatzis et al., 2001), STM1 is a putative GPCR from *S. pombe* (Chung et al., 2001), ERS1 complements a glycosylation mutant of yeast (Hardwick and Pelham, 1990), and MPDU1 (the human homologue of

CHO LEC35) is required for the use of Man/Glc-P-Dol in glycosylation (Anand et al., 2001). MPDU1/LEC35 and CTNS functional homologues have now been isolated from multiple species (Town et al., 1998; Cherqui et al., 2000; Anand et al., 2001; Yang et al., 2005). In this chapter, unless otherwise stated, MPDU1 and CTNS refer to the human forms.

For reference, I have split the family members into sub-groupings based on the phylogenetic tree. Group 1 contains *AtPQL1-3* and *STM1*, group 2 contains *AtPQL4*, *AtPQL6* and *MPDU1*, group 3 contains *CTNS* and *AtPQL5*, and group 4 contains *ERS1*.

As shown in figure 3-4b, the spacing of PQ-loops is conserved amongst all family members. Each has one PQ-Loop in each half of the protein. The spacing of transmembrane segments and the total number of amino acid residues varies, but is broadly conserved within sub-groups.

Alignment of full-length protein sequences (figure 3-5) shows that although the PQ-Loop is the best conserved part, there is some conservation amongst family members across the length of the protein sequence. Pairwise comparisons (table 3-2) indicate that sequence identity across the whole length of the protein can be as high as 74% for some pairings. The relationships between the PQL proteins within each sub-group are described below.

3.3.2.2 Sub-group 1: *AtPQL1*, *AtPQL2*, *AtPQL3* and *STM1*

The four proteins in sub-group 1 show a conserved arrangement of transmembrane domains and PQ-Loops. In each case, transmembrane segments 1-3 and 5-7 are closely spaced, with a large gap between transmembrane segments 3 and 4. The length of this gap varies between 50 residues for *STM1* to 119 residues for *AtPQL3*. *AtPQL1* and *AtPQL2* are particularly similar (57.1% identity). The most noticeable difference between the two is the length of the gap between transmembrane segments 3 and 4. For *AtPQL1* this gap is 59 residues whereas for *AtPQL2* it is 118. The varying length and sequence diversion in this region is illustrated by the alignment in figure 3-6a.

Table 3-2. Pairwise comparisons of PQL proteins. *AtPQL1-6*, MPDU1, CTNS, STM1 and ERS1 were aligned to each other in pairs. For each alignment the percentage identity (A), percentage similarity (B) and Blosum62 score (C) is given. Colouring is according to the Blosum62 score. Pairs with Blosum62 scores greater than 50 are coloured light blue and pairs with Blosum 62 scores greater than 200 are coloured dark blue.

		CTNS	MPDU1	STM1	ERS1	<i>AtPQL1</i>	<i>AtPQL2</i>	<i>AtPQL3</i>	<i>AtPQL4</i>	<i>AtPQL5</i>	<i>AtPQL6</i>
CTNS	A		16	15	20	15	12	5	17	30	16
	B		25	25	33	25	22	8	28	43	27
	C		69	98	287	89	67	59	98	564	94
MPDU1	A	16		9	8	16	11	13	29	18	30
	B	25		15	16	27	17	22	44	33	46
	C	69		52	46	47	42	39	290	64	293
STM1	A	15	9		16	25	22	24	18	18	18
	B	25	15		27	38	32	37	33	31	34
	C	98	52		47	259	253	279	88	109	76
ERS1	A	20	8	16		21	17	17	19	26	17
	B	33	16	27		36	25	29	32	44	27
	C	287	46	47		90	76	92	64	261	55
<i>AtPQL1</i>	A	15	16	25	21		57	36	18	18	19
	B	25	27	38	36		65	51	30	34	30
	C	89	47	259	90		1046	667	90	105	105
<i>AtPQL2</i>	A	12	11	22	17	57		39	11	6	16
	B	22	17	32	25	65		55	18	11	26
	C	67	42	253	76	1046		706	91	56	91
<i>AtPQL3</i>	A	5	13	24	17	36	39		19	19	19
	B	8	22	37	29	51	55		29	32	29
	C	59	39	279	92	667	706		72	100	73
<i>AtPQL4</i>	A	17	29	18	19	18	11	19		21	74
	B	28	44	33	32	30	18	29		34	87
	C	98	290	88	64	90	91	72		90	919
<i>AtPQL5</i>	A	30	18	18	26	18	6	19	21		18
	B	43	33	31	44	34	11	32	34		34
	C	564	64	109	261	105	56	100	90		50
<i>AtPQL6</i>	A	16	30	18	17	19	16	19	74	18	
	B	27	46	34	27	30	26	29	87	34	
	C	94	293	76	55	105	91	73	919	50	

3.3.2.3 Sub-group 2: *AtPQL4*, *AtPQL6* and *MPDU1*

AtPQL4 and *AtPQL6* are very closely related to each other (74% sequence identity) and are the closest *A. thaliana* relatives of *MPDU1* (29.3 and 30.4% sequence identity respectively). Blast searches using *MPDU1* as a query found no other *A. thaliana* homologues with significant similarity (see table 3-4). An alignment of the three sequences (figure 3-6b) shows that there is homology across the length of the sequence and that similarity is not confined to the PQ-loops.

In addition, the three proteins have a similar organisation of transmembrane domains and PQ-loops (figure 3-4b). In each case the transmembrane segments are evenly distributed across the protein with short hydrophilic segments at the N and C-terminals. *AtPQL4* and *AtPQL6* are particularly similar to each other in this respect. The predicted position of transmembrane segments varies by a maximum of just two residues. All three are of similar size, each containing between 234 and 247 amino acid residues.

3.3.2.4 Sub-group 3: *AtPQL5* and *CTNS*

AtPQL5 is the closest *A. thaliana* homologue of *CTNS* (see table 3-4). The two proteins share 29.9% amino acid identity. One of the most noticeable differences between the two is the presence of a long N-terminal hydrophilic tail in *CTNS* that is absent in *AtPQL5*. The number of residues before the first transmembrane domain in *CTNS* is 119 compared to just 13 in *AtPQL5*. None of the *AtPQL* proteins have such a long hydrophilic N-terminal.

A sequence alignment of *AtPQL5* and *CTNS* reveals considerable similarity in the PQ-loop domains and the region in between (see figure 3-6c) with more divergence at the N and C-termini.

A

<i>AtPQL2/1-376</i>	1	-----MFLHSSFRDGLSLSLGII SVI SWGVAEIPQIMTNYSKSTEGLSITFLTWTWIMGDI FN	58
<i>AtPQL1/1-288</i>	1	-----MIRDDL SLSLGII SVI SWSVAEIPQIMTNYNKSKI EGVSITFLTWTWMLGDI FN	53
<i>AtPQL3/1-374</i>	1	MVSLGYCLKEKKT CVRWVEI YFDDCLCNLNDVVSFALG IASLLCWGVAEIPQVI TNFRKTSSNGVSL SFL LAWVAGDI FN	80
<i>SpSTM1/1-271</i>	1	-----MSV IAPSFDMANI LTELSSFLGALSLGCWVLLI PQLLLENYKNQSGESI SDLFLI IWLIGDFFN	64
<i>AtPQL2/1-376</i>	59	LLGCLMEPATLPTQFY WALLYT VTTSVLYVQSI YYGHI YPRLKNRRDQMVEAERI SNII SDVKI PGRWRNSSD TTTTCGGQ	138
<i>AtPQL1/1-288</i>	54	VVGCLMEPASLPVQFY TAVLYTLATLVLYVQSI YYGHI YPRLMKNR	99
<i>AtPQL3/1-374</i>	81	LVGCLLEPATLPTQFY TALLYT VSTVVLVIQTI YYDYI YKLCRHRRTKICQKDEEDEERPL KP P KTMG	149
<i>SpSTM1/1-271</i>	65	VLGSI YGNVSS- TVLVLSFYI YVSDSTLMQIYYRWKAARRIASR	109
<i>AtPQL2/1-376</i>	139	TTPI TMIPGSQRTSFTGR ELFYT SARSLSSSHTPPAGSVLAQRMARGY SEPTLEEPLLPEDVTH PSTKSLLCVVS VFLFL	218
<i>AtPQL1/1-288</i>	100	-----RNHQMVDVEEPLLR EEA KR PSTKSLLCVVS VFLFL	134
<i>AtPQL3/1-374</i>	150	SAISI PPGSYKDSRR E FYT	205
<i>SpSTM1/1-271</i>	110	-----SARSLAGSGT PLRTSYFR VAKSGLRLWQSTMMVR	137
<i>AtPQL2/1-376</i>	219	GT FNLPNL LSESR TMLGEGDRVFVVR-- AARKLQV TSSNVAEHSGGESSRIGMFLGWAMAAI YMGGRLPQICLNMRRG	296
<i>AtPQL1/1-288</i>	135	GSF-- --NVLSGSR SMDLRGKDRVFVVGAGARKLLE VSSGNLGENN-- --IGMWLGWAMAAI YMGGRLPQICMNVRRG	206
<i>AtPQL3/1-374</i>	206	HRTKMRQCRRA GFGTFLAASASLPQA KSLAEKIAHASSRRLNERI VEHSSALGQWLWMAAI YMGGRIPQIWLNI KR	285
<i>SpSTM1/1-271</i>	138	-----LSTRQFQVMGC VIVSTI VGNLI I SSASSDKSDD-- --LNAWPFTAGCIS SVLYFCARIPQII KNHKAK	204
<i>AtPQL2/1-376</i>	297	HVEGLNPLMFFFA LVGNMTYVASI LVN-- SVEWLKLAPNLPWLVDAGGCVVLD FLILQFFHFRCKR-DKDTDKKKHETG	373
<i>AtPQL1/1-288</i>	207	NVEGLNPLMFFFA FIGNVTYVASI LVN-- SVEWSKIEPNLPWLVD SGGCAVLDFLILQFFYFHRCKVEADSVKKKHETG	284
<i>AtPQL3/1-374</i>	286	SVEGLNPLMFI FALVANATYVGSILVR-- TTEWONIKPNLPWLLDAI VCVVLD FLILQYIYYKYCRISKLESREEDAYG	363
<i>SpSTM1/1-271</i>	205	STEGLSI IFFVLASVGN TSYAFSILVFPASDYLNYTYA NLPWILGAFSTIFLDIYIF YQFI KYRNHY	271
<i>AtPQL2/1-376</i>	374	EAV-----	376
<i>AtPQL1/1-288</i>	285	EEAV-----	288
<i>AtPQL3/1-374</i>	364	DYVEASKTFVS	374
<i>SpSTM1/1-271</i>		-----	

B

<i>AtPQL6/1-234</i>	1	-----MDYLGIDMSCAIGSLRNGDFPEKDC LLPLISKLLGYCLVAASITVKLPQIMKI VQHKSVRGLSVVAFE	68
<i>AtPQL4/1-239</i>	1	-----MDYLGIDLSCAIGSLRNGEFPKDC LLPLISKLLGYFLVAASMTVKLPQIMKI VDNKSVKGLSVVAFE	68
<i>HsMPU1/1-247</i>	1	MAAEADGPKRL LVPILLPEKC YDQLFVQWDL LHVPC LKILL SKGLGLGI VAGSLLVKLPQVFKIRGAKSAEGLSLQSV	80
<i>AtPQL6/1-234</i>	69	LEVVGYSI SLAYCLH KGLPFSAFGE MAFLLI QALILVACIYYYSQPVVPTTWIRPLLYCAVAPT VLAGQINPTLFEALYA	148
<i>AtPQL4/1-239</i>	69	LEVIGYSI SLAYCLN KDLPFSAFGE LAFLLI QALILVACIYYFSQPLSVTTWVKAILYFAIAPT VFAKGLDPFLFEALYA	148
<i>HsMPU1/1-247</i>	81	LELVALTGTMYV SITNNPFSSVGEAL FLMLQTI TICFLVMHYRG-QTVKGVAFLAC YGLVLLVLLSPLTPLTVVTL LQA	159
<i>AtPQL6/1-234</i>	149	SQHAIFL FARLPQIWKNFKNKSTGELSFLTFFMNFAGSIVRVFTSLQE KAPISILTG FALGVVTNGSIL TQILLYS----	224
<i>AtPQL4/1-239</i>	149	SKHLIFLSARIPQIWKNFERNKSTGQLSFLTCLMNFAGALARVFTSIQE KAPLSMLLGIVLSIFTNGI IMSQILLYR SKGN	228
<i>HsMPU1/1-247</i>	160	SNVPADVVRLLQAATNYHNGYTQGLSAITVFL FGGSLARIFTSIQETG DPLMAGTFVVS LCNGLIAAQLLFYWN----	236
<i>AtPQL6/1-234</i>	225	KPAAAKEKKA-	234
<i>AtPQL4/1-239</i>	229	EDKL VKSKKIS	239
<i>HsMPU1/1-247</i>	237	AKPPH KQKKAQ	247

C

<i>AtPQL5/1-270</i>		-----	80
<i>HsCTNS/1-367</i>	1	MIRNWL TIFILFPLKL VEKCESSVSLTPPVVKLENGSSNTVSLTLRPP LNATLVI TFEITFRSKNITILELPDEVVVPP	80
<i>AtPQL5/1-270</i>	1	-----MASWNSIPLEI SYEIVGWI AFASWSISFY PQLILNFRRRSVVGLNF	46
<i>HsCTNS/1-367</i>	81	GVTNSSSFQVTSQNVGQLTVYLHGNH SNQTGPRIRFLVIRSSAISI INQVIGWI YFVAWSISFY PQVIMNWRKSVI GLSF	160
<i>AtPQL5/1-270</i>	47	DFVMLNLT KHSSYMI YNVCLYFSPVI QKQYFDTYGDKEMI PVAANDVAFS IHAVVM TAVTLFQIF IYERGPQKVSRLAIG	126
<i>HsCTNS/1-367</i>	161	DFVALNLTGFVAYS VFNIGLLWVPYI KEQFL LKYPNG-VNPNVNSDVF FSLHAVVLT LIIIVQCCLYERGGQRVSWPAIG	239
<i>AtPQL5/1-270</i>	127	IVVVVWGFAAICFFI ALP-THSWLWLSI FNSIQVMTCKYI PQAKMNFTRKSTVGSW SIGNILLDFTGGLANYLQMVIG	205
<i>HsCTNS/1-367</i>	240	FLVLAWLFAFVTMIVA AVGVITLQLFLCF SYIKLAVTLVKYFPQAYMNFY YKSTEGWSIGNVLLDFTGGFSFLQMFLO	319
<i>AtPQL5/1-270</i>	206	STDQNSWKNFYGNMGTLLSLI STFFDILFMFQHYVLYPEKKVSKSPETGEESNEP LIDSSHEHV	270
<i>HsCTNS/1-367</i>	320	SYNNQDWTLIFGDPTK FGLGVFSIVFDVVFIFQHFC L YRKR-----PGYDQLN----	367

Figure 3-6. Alignments of *AtPQL* proteins with their closest characterised non-plant homologue. The closest homologue of each *AtPQL* protein was determined using Blast (see table 3-4) and alignments were created using ClustalW software. *AtPQL1-3* are aligned to STM1 (A), *AtPQL4* and *AtPQL6* are aligned to MPDU1 (B), and *AtPQL5* is aligned to CTNS (C). Alignments were coloured as described for figure 3-5.

3.3.3 Identification of potentially important residues in the PQ-loop

The function of the PQ-loop is unclear. However, a number of naturally occurring point mutations in the PQ-loops of the human PQL-proteins, CTNS and MPDU1 have been reported to lead to loss of function phenotypes and site-directed mutagenesis of the PQ-loops of CTNS and STM1 give indications of the importance of some key residues.

Point-mutations in CTNS and MPDU1: In total 16 different mutations have been identified in the PQ-Loops of CTNS leading to single amino acid substitutions and clinical phenotypes (Shotelersuk et al., 1998; Town et al., 1998; Attard et al., 1999; McGowan-Jordan et al., 1999; Thoene et al., 1999; Kalatzis et al., 2002; Kalatzis et al., 2004). Of these, eight affect the first PQ-loop and eight the second. None of these mutations were found to affect the targeting of the protein to the lysosomes but most reduced cystine transport activity (Kalatzis et al., 2004).

Most of these mutations affect residues that do not show a high level of conservation throughout the PQL family. However, five are in highly conserved residues (see highlighted residues in table 3-3). Of these five mutations, four have been reported to almost completely abolish cystine transport whilst the other, S298N had only a modest effect (cystine transport 77% of wild type; Kalatzis et al., 2004). Interestingly two of these mutations, G169D in the first PQ-loop and G308R in the second, correspond to each other when the two PQ loops are aligned.

For the most part, these mutated residues are conserved in both PQ-loops of *AtPQL1-6*. Apart from one substitution, all five residues are the same in both PQ-loops of *AtPQL1* and *AtPQL2*, whilst *AtPQL4* and *AtPQL6* differ by one change each. *AtPQL3* and *AtPQL5* are more diverse in these residues differing by four substitutions each.

Mutations have also been reported in the PQ-loop of the human protein MPDU1 (Kranz et al., 2001; Schenk et al., 2001b). Two mutant forms have been reported (see table 3-3). As highlighted in figure 3-2, they occur next to each other in the first PQ-loop. One of these, L74S, is in a strongly conserved residue and aligns to a known mutation in the first PQ-loop of human CTNS (L158P). Both of the mutated residues are conserved in the closest *A. thaliana* relatives of MPDU1 – *AtPQL4* and *AtPQL6*.

Site directed mutagenesis of CTNS and STM1: Site directed mutagenesis has identified further potentially important regions in the PQ-loop. A YFPQA motif, identified in the second PQ-loop of CTNS was found to be important, along with a C-terminal signal, for correct targeting to the lysosomes (Cherqui et al., 2001). Figure 3-7a shows an alignment of this motif with the corresponding residues of *AtPQL1-6*. Apart from the PQ-dipeptide, this motif is generally not conserved between *AtPQL1-6* and CTNS. *AtPQL5* is the only *A. thaliana* protein that shows conservation of the other residues in this motif. The tyrosine and alanine at positions one and five are conserved in *AtPQL5*, but not the phenylalanine at position two. Although the residues in these positions are not conserved between *AtPQL1-6* and CTNS, there does appear to be some conservation within the *A. thaliana* PQL family. With the exception of *AtPQL5* the motif is always R-I/L-PQL.

Site-directed mutagenesis of the second PQ-Loop of STM1, a predicted GPCR from *S. pombe*, also identified functionally important residues (Chung et al., 2001). A putative binding site for interaction with the G α subunit, GPA2 was identified through homology with other GPCRs from yeast (Chung et al., 2001). Mutation of a lysine residue at position four of this motif was found to abolish the phenotype of an STM1 over-expresser. Figure 3-7b shows an alignment of the putative binding site motif in STM1, other yeast GPCRs and the corresponding amino acids in *AtPQL1-6*. As can be seen, the motif is much better conserved in the yeast proteins than *AtPQL1-6*. The only residue matching the putative binding site sequence in all six *A. thaliana* proteins was the glutamine at position eight in the motif. All but *AtPQL5* also had an arginine or lysine at position one. None matched the binding-site sequence at positions two and three, and only *AtPQL4* and *AtPQL6* shared the crucial lysine at position four.

Table 3-3. Point mutations in CTNS and MPDU1. Mutations reported to lead to clinical phenotypes are indicated along with the equivalent residues in other PQL proteins. Mutations in highly conserved residues are highlighted yellow (conservation score ≥ 5 , see figure 3-1). The equivalent residue in other PQ-Loops is coloured dark blue if it is identical to the mutated residue or light blue if it is similar (i.e. Gonnet Pam250 $x > 0.5$). *Transport activity of mutant forms of CTNS is given as percentage of wild type as determined by Kalatzis et al. (2004). Negative values indicate less uptake than an untransformed control (see section 1.2).

PQ-Loop	Mutation	Ref	Alignment Score	Transport Activity*	Conservation of residue in other PQ-Loops																			
					AtPQL1		AtPQL2		AtPQL3		AtPQL4		AtPQL5		AtPQL6		STM1		MPDU1		CTNS		ERS1	
					1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
CTNS (1st)	Q128X	1	N/A	N/A	L	-	L	M	F	Q	K	-	E	S	K	-	S	-	K	-	Q	F	D	-
	W138X	1	4	N/A	W	Y	W	Y	W	Y	S	F	W	T	S	F	W	Y	S	V	W	T	W	S
	S141F	2	4	2.0 \pm 5.3	S	M	G	M	G	M	M	L	S	C	I	L	V	F	L	V	S	L	S	L
	L158P	3	6	1.4 \pm 4.2	V	L	L	L	V	L	L	L	W	L	L	I	L	L	L	L	W	I	F	F
	G169D	2	5	-0.82 \pm 1.4	G	G	G	G	G	A	G	G	K	G	G	G	G	A	G	G	G	G	G	G
	Y173X	1	2	N/A	N	Y	N	Y	N	Y	S	R	Y	N	S	R	N	Y	T	R	Y	S	L	S
	N177T	4	4	0.70 \pm 4.1	C	I	C	I	C	I	D	S	N	M	C	S	S	I	N	S	N	M	I	L
	W182R	5	1	34 \pm 5.9	A	S	A	S	A	T	G	A	F	I	G	A	V	A	G	G	W	Y	Y	S
CTNS (2nd)	K280R	6	3	-0.68 \pm 0.9	A	G	A	G	A	G	V	A	S	K	V	A	L	A	V	G	S	K	S	K
	N288K	4	5	1.6 \pm 1.2	N	N	N	N	N	N	I	N	N	N	I	N	N	N	I	N	N	N	N	N
	Y290X	5	2	N/A	Q	R	E	R	T	R	N	N	R	R	H	N	N	A	A	N	R	Y	H	R
	S298N	5	6	77 \pm 21	S	N	S	N	S	N	S	S	N	S	S	S	S	S	S	S	S	S	S	P
	D305Y	7	2	-4.8 \pm 1.7	W	A	W	A	W	A	E	N	N	D	E	N	W	A	E	L	N	D	N	D
	G308R	5	5	-5.3 \pm 8.5	G	G	G	G	G	A	G	G	K	G	G	G	G	A	G	G	G	G	G	G
	S310Q	5	4	N/A	I	V	I	A	I	A	T	L	S	L	T	I	F	T	T	L	V	S	S	I
	N323K	6	1	0.14 \pm 0.8	L	E	L	E	L	E	P	L	P	Q	P	I	S	D	P	D	P	N	W	D
MPDU1 (1st)	G73E	8	4	N/A	G	G	G	G	G	G	Q	G	G	G	E	S	G	G	Q	G	G	A	C	
	L74S	9	6	N/A	V	L	L	L	V	L	L	L	L	W	L	L	I	L	L	L	W	I	F	

References for table 3-3:

- | | | |
|--------------------------------|-----------------------------|-------------------------|
| 1 Town et al. (1998) | 4 Kalatzis et al. (2002) | 7 Attard et al. (1999) |
| 2 McGowan-Jordan et al. (1999) | 5 Shotelersuk et al. (1998) | 8 Schenk et al. (2001b) |
| 3 Kalatzis et al. (2004) | 6 Thoene et al. (1999) | 9 Kranz et al. (2001) |

A

CTNS YFPQA
AtPQL1 RLPQI
AtPQL2 RLPQI
AtPQL3 RLPQI
AtPQL4 RLPQI
AtPQL5 YIPQA
AtPQL6 RLPQI

B

Ga binding site KK^{*}K LG
RR^{*}R VQ

GPR1-3 KR I KAQ IG
GPR1-5 KKRR AQ IQ
STE2 RSRR F LG
STE3 KKR KD VR
MAM2 L I R KK IG
MAP3 RRQR E LQ
STM1 KHN KI IQ
AtPQL1 RVNMC IQ
AtPQL2 RMNLC IQ
AtPQL3 K I N LW IQ
AtPQL4 R FN KW IQ
AtPQL5 T F N MK AQ
AtPQL6 K FN KW IQ

Figure 3-7. Alignment of known functional residues in CTNS and STM1 with equivalent residues in AtPQL1-6. Alignments were created manually following initial alignment of full length sequences (see figure 3-2) **A.** Lysosomal targeting signal from CTNS. **B.** Ga binding site from STM1. Also aligned are other known yeast GPCRs and the consensus sequence presented by Chung et al., (2001). The lysine residue essential for function is marked with an asterisk. In both cases matching residues are highlighted.

3.3.4 Similarity to non-PQL proteins

Pfam searches were used to look for other known domains in the *AtPQL* sequences (Finn et al., 2006). The PQ-loop is the only domain common to all six *AtPQL* proteins. However, other domains are recognised as having weak homology to the *AtPQL* sequences. *AtPQL6* shares some homology to the tellurite-resistance/dicarboxylate transporter (TDT) domain (e-value 0.045; PF03595) and both *AtPQL4* and *AtPQL6* have weak homology to the 7TMR-DISM 7TM domain (e-value 0.27 for *AtPQL4* and 0.11 for *AtPQL6*; PF07695).

To find other similar proteins, Blast searches were performed using the full-length amino acid sequence of each PQL protein as a query against the *A. thaliana* reference sequence database (Altschul et al., 1997). All hits with an e-value below 1 are given in table 3-4. When using *AtPQL1-6* as a query, most hits were for other *AtPQL* proteins. The exceptions were a protein of unknown function (At5g27010) showing weak similarity to *AtPQL1* (e-value 0.58) and a U3 ribonucleoprotein (UTP) complex protein (At5g08600) showing similarity to *AtPQL3* (e-value 0.076). The Blast alignment of *AtPQL3* and the UTP complex protein starts at the end of the second cytoplasmic loop of *AtPQL3* and spans the fourth transmembrane domain and third non-cytoplasmic loop. This region corresponds to a region close to the centre of the UTP complex protein. UTP proteins are known to form subunits of the U3 ribonucleoprotein complex and are required for the synthesis of 18S ribosomal RNA (Dragon et al., 2002).

3.3.4.1 *AtPQL4* and *AtPQL6* and the 7TMR-DISM 7M domain

The 7TMR-DISM 7TM domain is the transmembrane portion of the 7TMR-DISM (seven transmembrane receptors with diverse intracellular signalling modules) family. This family was identified in a database search for candidate prokaryotic receptor proteins with multiple transmembrane segments and large extracellular domains (Anantharaman and Aravind, 2003). They are similar to the PQL family in that all members possess seven transmembrane-spanning domains with the N-terminal facing outwards. However, whereas PQL proteins usually have short non-membrane N and C-

terminals, the 7TMR-DISM family has large hydrophilic domains at both ends. The N-terminal region is predicted to be involved in ligand-binding, whereas the C-terminal is typically a signalling domain such as a PP2C phosphatase or histidine kinase catalytic domain. The core transmembrane domain is usually 250 to 300 amino acids in length, similar to *AtPQL4* and *AtPQL6*.

Figure 3-8 shows an alignment of the 7TMR-DISM 7TM domain of selected 7TMR-DISM family proteins with the full-length sequences of *AtPQL4* and *AtPQL6*. The homology is spread across the length of the sequences without any particular regions showing greater similarity.

3.3.4.2 *AtPQL6* and the TDT domain

Tellurite-resistance/dicarboxylate transporter (TDT) domain containing proteins form a diverse family of integral membrane proteins (Saier et al., 1999). They are present in both prokaryotes and eukaryotes and are mostly of unknown function. Two members of the family have been characterised. The first, TehA is a proton-coupled ethidium/proflavin efflux transporter from *E. coli* (Turner et al., 1997). The second, Mae1 is a proton-coupled malate/succinate/malonic acid uptake transporter from *S. pombe* (Grobler et al., 1995).

Full-length alignments of *AtPQL6* and selected members of the TDT family are shown in figure 3-9a. The similarity is not confined to any particular part of *AtPQL6*, but spread evenly across the length of the sequence. Figure 3-9b shows two short signature sequences which have been identified for the TDT family (Saier et al., 1999). As can be seen, with the exception of four residues in the first and six in the second, these signatures are not conserved in *AtPQL6*.

3.3.4.3 *The AtPQL family and the TRP ion channel family*

In order to detect more distant relationships between the *AtPQL* family and other protein families, the full-length alignment of *AtPQL1-6* was queried against the Pfam database in a COMPASS search (Sadreyev and Grishin, 2003). Querying the alignment

rather than individual members may uncover more distant relationships because general features of different families may be similar even when homology between individual members is low. Significant homology was found to the fungal transient receptor potential (TRP) ion channel family (PF06011; e-value 1.59×10^{-7}).

An alignment of the consensus sequences for *AtPQL1-6* with the Pfam consensus sequence for the TRP family is shown in figure 3-10. As can be seen there is similarity throughout, particularly in the PQ-loops and a region close to the centre of the *AtPQL* sequence. This family of TRP channels (PF06011) are distinct from the mammalian TRP II channels (PF08344). One member of this family, *S. pombe* PKD2, is proposed to be a golgi localised calcium channel involved in cellular calcium homeostasis and the control of cell wall morphology (Palmer et al., 2005).

```

AtPQL6/1-234      MDYLGIDMSCAIGSLRNGDFPEKDCLLPLISKLLGYCLVAASITVKLPQIMKIVQHKSVRGLS--VVAFELEVVGYTISLAY
AtPQL4/1-239      MDYLGIDLSCAIGSLRNGEFPKDCCLPLISKLLGYFLVAASMTVKLPQIMKIVDNKSVKGLS--VVAFELEVIGYITISLAY
Pa_Q9HUV7_187-390 -----AYAFGMLLGG-----MLLLMYNLIRFAYSRSSSLWLAHVHAALAVCAAANLG--LVAFWLPGLKFNQSLTA
Pa_Q9HYE4_180-383 -----LMGFNGAFYG-----VLFGLFCYNLRFVSLREATYAWYLLYNLSGLFSASFDG--LLFKLLPGHVALESAGI
Td_Q3SFG4_182-387 -----QELSYGAVYG-----FLFALLVYNAILYASLGGARYISYAFYLA MFMLMNIAYTG--HGYAWLPPESTVWQQWS
Bh_Q9KBB5_85-286 -----TSAQALMVAA-----IMFIHGIIYACILFFIRPKKIELLYFMVAVLLSAFSLVLSDERLLQWVPLNYEWFIKLA
Le_Q8F2U2_257-456 -----IQIPNGLIRKD-----EVNII LVI EVKKYFQKEIGIEQDKTAIGD SLLIQKELLE T--EYIKILLMIYTTAGV

AtPQL6/1-234      CLHKGLPFSAFGEMAFLLIQALILVACIYYYSQVPV-VTTWIR--PLLYCAVAPTVLAGQINPTLFEALYASQHAIFLFARL
AtPQL4/1-239      CLNKDLPFSAFGELAFLLIQALILVACIYYFSQRLS-VTTWVK--AILYFAIAPT VFAGKIDPFLFEALYASKHLIFLSARI
Pa_Q9HUV7_187-390 DLG-----ALGAAVSLWFAQSFRRGTAE SRLNRLLQGEALLILAVG-AIIAFTQQLWFSWLIIYLLVILSSLSVPL
Pa_Q9HYE4_180-383 YL-----LMYLSCLVSIQFSRGYLYTRDFPRLD RFLRG-LLLACVLLASEPLVGLRAWNVLASLTVMVLSL SLL
Td_Q3SFG4_182-387 NP-----VLMYLYGAAGLLFAVRFDLGVFFPRVRRAVEVYSAAGTALVAAVAAGRQVAALVVSFSFVFLFTGLMLA
Bh_Q9KBB5_85-286 YLS-----YTGLSLFFLLFVRYLFSNYKRNQSVRWISG-FCALYALFILIAPATTIREWTIMLLFVLAIPFTTII L
Le_Q8F2U2_257-456 Y-----FLLYLRRRADRENLYYGLFTILLVLYQLRNQIKYDLGIEFIYMKKLEYMVLATLVPLFANFIRLYF

AtPQL6/1-234      PQIWKNFKNKSTGELSFLTFFMNFAG--SIVRVFTSLQEKAPI SILTGFA LGVVTVNGSILTQILLYS---KPAAAKEKKA-
AtPQL4/1-239      PQIWKNFNRKSTGQLSFLTCLMNFEG--ALARVFTSIQEKAPLSMLLGIVLSIFTNGIIMSQILLYRSKGNEDKL VKSKKIS
Pa_Q9HUV7_187-390 IAAWHYRGYQPARLIVAGMIYENAGFMVFLPVLFGTKQLDPGWLVLGVSFATLAGLVLSVSLTER-----
Pa_Q9HYE4_180-383 AGVHVWRQGLRYGLYYILAWGALLLSFLTVAASLGFLFG-LFGSSVVKIGMTVELVTLISGLADR-----
Td_Q3SFG4_182-387 LGLIAYRAGQKPARYFLIAALAAVMGAALTTLSTWGFVPHN-AWTFRAVEVGMLFDATLLALALAYQFR-----
Bh_Q9KBB5_85-286 LVVQVILKKERDSLFLLLAAISVASS--ILWSSFTGRVKSFEWGLSQIDYSFYPVDLTAFLC-----
Le_Q8F2U2_257-456 KYPKTTILKVL DGLYVSFAIYFFSN--NVEHYNI FINKNLVQYGWILYLGITLNHLIRRVIAKDRD-----

```

Figure 3-8. Similarity between the 7TMR-DISM 7TM domain and *AtPQL4* and *AtPQL6*. Alignment of the 7TM domains of a variety of 7TMR_DISM proteins with the full-length sequences of *AtPQL4* and *AtPQL6*. Colouring is as described for figure 3-5. 7TMR-DISM sequences are from *Pseudomonas aeruginosa* (*Pa*), *Bacillus halodurans* (*Bh*) and *Leptospira interrogans* (*Le*).

Table 3-4. Blast searches using PQL proteins as queries. Hits are shown from Blast searches using the full length protein sequences of CTNS, STM1 MPDU1, ERS1 and *AtPQL1-6* as queries. The BlastP program (cut-off value of 1) and the *A. thaliana* ref-seq database were selected. The alignment length, position, and e-value are given for each hit. The percentage identity and similarity within the alignment is also included. Hits for other *AtPQL* proteins when using *AtPQL* proteins as queries have been omitted.

Query	Hit	AGI Accession	Alignment Length	% Identity	% Positives	Aligned Region (query)		Aligned Region (hit)		e- value
						Start	End	Start	End	
CTNS	<i>AtPQL5</i>	AT5G40670	242	46.7	68.2	120	360	6	246	3.00E-59
	<i>AtPQL1</i>	AT4G20100	56	33.9	60.7	130	184	10	61	0.025
	SOS1	AT2G01980	98	29.6	44.9	216	311	203	294	0.043
	<i>AtPQL3</i>	AT4G36850	104	26.9	46.2	126	227	33	116	0.074
	<i>AtPQL2</i>	AT2G41050	70	30.0	52.9	116	184	1	66	0.13
	KUP6	AT1G70300	178	23.6	39.9	111	287	226	398	0.48
	KUP2	AT2G40540	171	22.2	36.8	130	287	236	394	0.63
	<i>AtPQL4</i>	AT5G59470	246	22.8	37.0	123	366	27	232	0.63
	repair endonuclease	AT5G39770	60	26.7	53.3	15	73	410	469	0.82
MPDU1	<i>AtPQL4</i>	AT5G59470	203	36.0	54.2	37	234	25	223	8.00E-27
	<i>AtPQL6</i>	AT4G07390	216	35.7	53.7	37	246	25	234	2.00E-26
ERS1	<i>AtPQL5</i>	AT5G40670	248	29.8	50.4	6	246	14	241	6.00E-21
	<i>AtPQL3</i>	AT4G36850	57	31.6	59.7	5	58	31	87	0.002
	<i>AtPQL1</i>	AT4G20100	51	33.3	56.9	8	58	10	60	0.005
	<i>AtPQL2</i>	AT2G41050	54	33.3	55.6	8	61	15	62	0.035
	FRO7	AT5G49740	172	22.7	43.6	7	164	38	183	0.86
STM1	<i>AtPQL1</i>	AT4G20100	274	27.7	46.0	13	267	2	267	2.00E-19
	<i>AtPQL3</i>	AT4G36850	87	40.2	58.6	182	268	263	347	4.00E-10
	<i>AtPQL3</i>	AT4G36850	93	35.5	54.8	10	98	26	115	6.00E-10
	<i>AtPQL3</i>	AT4G36850	96	30.2	50.0	171	265	30	118	0.013
	<i>AtPQL3</i>	AT4G36850	104	26.0	48.1	17	112	258	360	0.14
	<i>AtPQL2</i>	AT2G41050	83	39.8	56.6	17	98	11	93	6.00E-10
	<i>AtPQL2</i>	AT2G41050	108	32.4	50.0	161	268	253	358	3.00E-08
	<i>AtPQL2</i>	AT2G41050	53	41.5	56.6	182	234	16	68	7.00E-04
	<i>AtPQL6</i>	AT4G07390	102	24.5	44.1	4	105	129	230	0.063
	<i>AtPQL6</i>	AT4G07390	42	40.5	57.1	188	229	153	194	0.31
	<i>AtPQL6</i>	AT4G07390	46	39.1	56.5	184	222	30	75	0.92
	Nodulin/MtN3 family	AT5G50800	98	25.5	42.9	175	269	6	95	0.11
	<i>AtPQL4</i>	AT5G59470	83	30.1	48.2	187	269	152	225	0.11
<i>AtPQL1</i>	Unknown	AT5G27010	99	24.2	45.5	38	134	708	804	0.58
<i>AtPQL3</i>	Utp family	AT5G08600	79	29.1	44.3	183	261	304	374	0.076

3.3.5 Genomic organisation of *AtPQL1-6*

The *AtPQL* genes are distributed throughout the genome. *AtPQL1*, *AtPQL3* and *AtPQL6* are on chromosome 4, *AtPQL4* and *AtPQL5* are on chromosome 5 and *AtPQL2* is on chromosome 2. However, no two *AtPQL* genes occur in any of the corresponding duplicated regions known to be present in the *A. thaliana* genome (Simillion et al., 2002).

To examine possible inter-species *AtPQL* gene duplication, homologous regions of *A. thaliana*, *P. trichocarpa*, *O. sativa* and *S. bicolor* were examined using the plant genome duplication database (Wang et al., 2006). Each locus was used as a query to search the database and the results are summarised in table 3-5. *AtPQL2*, *AtPQL3* and *AtPQL5*, were all found to have *P. trichocarpa* (Poplar) homologues that occur in homologous regions of the *A. thaliana* and *P. trichocarpa* genomes. *AtPQL2* appears to be duplicated twice as part of two small homologous regions, whereas *AtPQL3* and *AtPQL5* are each duplicated once as part of large homologous regions.

3.3.6 Intron-exon organisation of *AtPQL1-6*

As can be seen in figure 3-11, *AtPQL1-6* do not share a common pattern of introns and exons. The intron number varies from zero in *AtPQL1* to eleven in *AtPQL2*. No evidence could be found of any conservation of intron placement even amongst the most closely related family members.

Table 3-5. Homologous *PQL* genes in duplicated regions of the *A. thaliana* and *P. trichocarpa* genomes. *AtPQL1-6* were used as queries to search the plant genome duplication database. The number of homologues in duplicated blocks is given along with the locus and chromosome of predicted duplicated genes.

Gene	Accession	No. of homologues in duplicated regions in <i>P. trichocarpa</i>	Locus	Chromosome
<i>AtPQL1</i>	At4g20100	0		
<i>AtPQL2</i>	At2g41050	2	TOP06.212 TOP16.0168	6 16
<i>AtPQL3</i>	At4g36850	1	TOP07.0974	17
<i>AtPQL4</i>	At5g59470	0		
<i>AtPQL5</i>	At5g40670	1	TOP17.0203	17
<i>AtPQL6</i>	At4g07390	0		

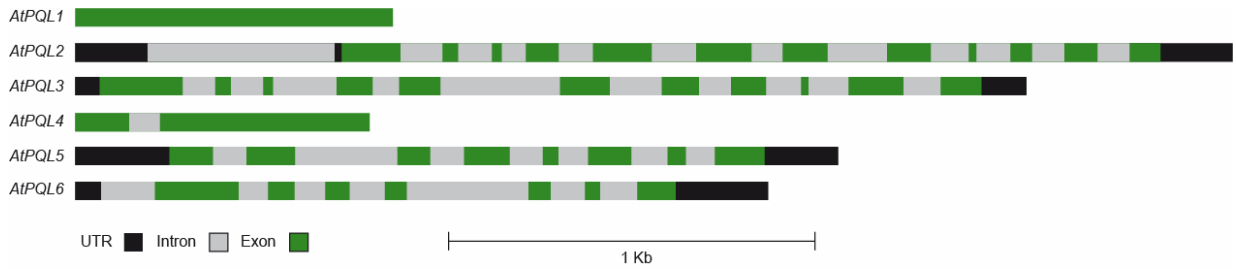


Figure 3-11. Intron-exon organisation of *AtPQL1-6*. Introns (grey), exons (green) and untranslated regions (black) where known, are drawn to scale for each gene.

Table 3-6. Genetic characteristics of *AtPQL1-6*. For each gene, the number of introns and the total length of the UTR and open reading frame are given.

AGI Accession	Gene	Total ORF Sequence (bp)	no. of introns	5' UTR (bp)	3' UTR (bp)
At4g20100	<i>AtPQL1</i>	867	0	N/A	N/A
At2g41050	<i>AtPQL2</i>	2235	11	216	197
At4g36850	<i>AtPQL3</i>	2407	10	66	124
At5g59470	<i>AtPQL4</i>	804	1	N/A	N/A
At5g40670	<i>AtPQL5</i>	1623	7	257	202
At4g07390	<i>AtPQL6</i>	1567	7	70	254

3.3.7 Promoter sequences of *AtPQL1-6*

To examine potential regulatory motifs in the promoter regions of *AtPQL1-6*, the Athena program was used (O'Connor et al., 2005). The region immediately 5' of the transcription start site (1 kb or up to the adjacent gene) was queried against a database of known cis-acting regulatory elements (see table 3-7). Although several known motifs were found in the promoter region of each gene, there was no clear pattern of significantly over-represented motifs and few indications for regulation. However, it is interesting that the upstream region of *AtPQL6* contains both an evening element promoter motif and a CCA binding box, both of which have been implicated in diurnal regulation of transcription (Wang et al., 1997; Harmer et al., 2000).

Table 3-7. Known transcription factor motifs upstream of *AtPQL1-6*. The genomic region 1 kb upstream of the transcription start site of each gene was scanned for known motifs. The strand column indicates whether the element is on the sense or antisense strand relative to the gene.

Gene	Strand	Promoter Motifs
<i>AtPQL1</i>	Sense	GAREAT, Hexamer, MYB1AT, T-box, TATA-box, W-box
	Antisense	BoxII, GAREAT, Hexamer, MYB, MYB4, TATA-box, TELO-box, W-box
<i>AtPQL2</i>	Sense	ATHB6, AtMYC2 BS in RD22, CARGCW8GAT, GAREAT, Hexamer, MYB4, TATA-box, W-box
	Antisense	CARGCW8GAT, DRE core, MYCATERD1
<i>AtPQL3</i>	Sense	GAREAT, MYB, MYB1AT, MYB4, T-box, TATA-box
	Antisense	ATHB2, GAREAT, MYB1AT, T-box, TATA-box
<i>AtPQL4</i>	Sense	ARF, GAREAT, LEAFYATAG, MYB, MYB1, MYB1AT, MYB2AT, MYB4
	Antisense	MYB1AT, TATA-box, T-box, W-box
<i>AtPQL5</i>	Sense	AGCBOXNPGLB, CARGCW8GAT, GCC-box, MYB1AT, MYB4
	Antisense	BoxII, CARGCW8GAT, DRE core, DREB1A/CBF3, GAREAT, GCC-box, MYB1AT, MYB4, TELO-box, W-box
<i>AtPQL6</i>	Sense	AtMYC2 BS in RD22, CARGCW8GAT, MYB1LEPR, MYCATERD1, TATA-box
	Antisense	ABRE, ABRE-like, ACGTA BRE, ATHB2, AtMYC2 BS in RD22, CARGCW8GAT, CCA1, EveningElement, GAREAT, MYB3, MYB4, MYCATERD1, TATA-box, W-box

3.4 Discussion

Six members of the PQL family have been identified in *A. thaliana*. Four more than were identified in previous studies (Zhai et al., 2001). None of these proteins have previously been characterised. In addition, sub-groupings of the PQL family have been identified and the similarities between these sub-groupings and non-plant PQL homologues have been described. Furthermore, novel relationships between PQL proteins and other protein families have been discovered.

3.4.1 Functional relevance of the PQ-loop

The relevance of the PQ-loop is unclear. Given the highly conserved hydropathy of the domain (see figures 3-2b), one possibility is that it simply represents an efficient way for proteins to form transmembrane domains.

However, some of the mutations affecting the PQ-loop in human CTNS suggest that certain residues within the PQ-loops may play additional roles. Kalatzis et al. (2004) found that several point mutations in the PQ-loop eliminate cystine transport activity. Although this could be a result of misfolding they do not lead to degradation or mislocalisation of the target protein, arguing instead for direct functional involvement in transport activity. In addition, site-directed mutagenesis of residues within the PQ-loops of CTNS and STM1 have uncovered motifs essential for trafficking and binding to interacting proteins respectively (Cherqui et al., 2001; Chung et al., 2001). Nevertheless, given the seemingly diverse functions of known PQL proteins, it is difficult to predict a common role played by the PQ-loop (see chapter six).

The best-conserved part of the PQ-loop is the PQ-dipeptide that gives the domain its name. The significance of this dipeptide is unclear but interestingly it always occurs close to the border between transmembrane and non-transmembrane regions. This may be significant as proline, unique in that it binds the protein backbone twice, frequently occurs where the polypeptide chain sharply changes direction (Betts and Russell, 2003).

3.4.2 Functional relevance of conserved topology

All six *AtPQL* proteins and all other PQL family proteins examined so far are predicted to have seven transmembrane domains with the N-terminal facing the non-cytoplasmic side of the membrane (Hardwick and Pelham, 1990; Town et al., 1998; Anand et al., 2001; Chung et al., 2001). This topology is often associated with receptors, particularly G-protein coupled receptors, and homology searches often aim to identify potential receptor proteins based on this topology (Anantharaman and Aravind, 2003; Moriyama et al., 2006).

However, the non-cytoplasmic N-terminal domains in *AtPQL1-6* are all relatively short and none contain any domains predicted to be involved in ligand binding. Furthermore, no cytoplasmic domains that may play a role in signal transduction could be identified. This is in contrast to other receptor proteins such as the 7TMR-DISM family (see section 3.3.4.1). In addition, *AtPQL1-6* differ in essential residues from the putative GPCR and PQL protein STM1, making a common function unlikely (see section 3.4.4.1).

3.4.3 Intron-exon and genomic organisation

There is no evidence of conservation of intron-exon structure amongst *AtPQL* genes. Interestingly *AtPQL1* has no introns, yet its closest homologues *AtPQL2* and *AtPQL3* have 11 and 10 introns each. One hypothesis for the origin of intron-less genes is that they arose from mRNA, reverse transcribed into cDNA and then recombined back into the genome (Liaud et al., 1995; Frugoli et al., 1998). This seems a plausible hypothesis for *AtPQL1*. Similarly *AtPQL4* has just one intron whereas its closest homologue *AtPQL6* has seven. Perhaps *AtPQL4* arose as a result of *AtPQL6* cDNA being reinserted into the genome and a single intron appearing later. This hypothesis would be one explanation as to why these gene pairings are not in corresponding duplicated regions of the genome. Furthermore, the lack of *AtPQL1* and *AtPQL4* homologues in duplicated regions between *A. thaliana* and *P. trichocarpa* support the hypothesis that these genes arose relatively recently.

3.4.2 Specific features of individual genes and sub-groupings

In addition to features common to the whole family, sequence analysis has uncovered features that may have functional implications for individual proteins. I will now discuss specific findings for each sub-group and pay particular attention to the similarity with non-plant PQL proteins.

3.4.4.1 Functional implications for sub-group 1: *AtPQL1-3*

AtPQL1-3 show homology to the *S. pombe* PQL protein, STM1. STM1 is a putative GPCR involved in nutrient sensing and cell-cycle regulation (Chung et al., 2001). Crucial for the function of STM1 is the $G\alpha$ binding site. The putative interaction motif consists of 7-8 residues and falls in the second PQ-loop of STM1. A lysine residue (position 199) within this motif has been shown to be essential (Chung et al., 2001).

These residues are conserved throughout several GPCRs (figure 3-7b) but not in *AtPQL1-3*. In particular, the lysine-199 is absent in *AtPQL1-3* arguing against $G\alpha$ binding activity. In addition, no significant homology could be found between any other GPCRs and the *AtPQL* proteins.

3.4.4.2 Functional implications for sub-group 2: *AtPQL4* and *AtPQL6*

AtPQL4 and *AtPQL6* are the closest *A. thaliana* homologues to MPDU1 sharing 29.3 and 30.4% identity respectively. In addition, two residues known to be essential for MPDU1 function are identical in *AtPQL4* and *AtPQL6* (see table 3-3). Given this and other data presented in this thesis, functional homology between these proteins and MPDU1 is predicted (see chapter six). MPDU1 is required for the use of Man/Glc-P-Dol as sugar donors during protein glycosylation in the ER lumen.

Interestingly, *AtPQL4* and *AtPQL6* also show weak homology to the 7TMR-DISM 7TM domain. Other proteins that show weak homology to this domain include ALG9, a yeast ER localised mannosyltransferase (Burda et al., 1996). ALG9 uses Man-P-Dol as a sugar donor to add the seventh and ninth Man residues to the LLO-precursor

during N-glycosylation (see chapter 1, figure 1-2, steps 12 and 14). If *AtPQL4* and *AtPQL6* are involved in Man-P-Dol utilisation, the presence of the 7TMR-DISM 7TM domain in these three proteins may reflect a common membrane organisation optimised for localisation to the ER membrane and/or utilisation of Man-P-Dol.

Interestingly the promoter region of *AtPQL6* has both a CCA1 binding site and an evening element motif. Both are connected to diurnal regulation. The function of the evening element motif is unclear but it has been identified as being particularly frequent amongst promoters of genes with peak expression at the end of the day (Harmer et al., 2000). The CCA1 binding site motif was first identified in the promoter of the light-harvesting chlorophyll a/b gene, *LHCBI*3* and found to be essential for its transcriptional regulation by red light (Wang et al., 1997). Expression of *CCA1* oscillates according to circadian rhythms and peaks at dawn (Wang and Tobin, 1998). It has been implicated as a key regulator of the circadian clock and is thought to act by repressing transcription in the morning (Alabadi et al., 2001). The combination of both an evening element and a CCA1 binding site suggests peak expression at the end of day.

3.4.4.3 Functional implications for sub-group 3: *AtPQL5*

AtPQL5 is the most similar *A. thaliana* protein to CTNS, a human lysosomal cystine transporter (Kalatzis et al. 2001). One key difference between CTNS and *AtPQL5* is a long N-terminal tail present in CTNS but absent in *AtPQL5*. This long N-terminal contains seven potential N-glycosylation sites and it has been proposed that it is highly glycosylated to protect CTNS from lysosomal protein hydrolysis (Cherqui et al., 2001). Assuming a CTNS-like protein in plants would localise to the vacuole (see section 1.2.2), the lack of a protective N-terminal tail in *AtPQL5* may reflect differences between the lysosomal and vacuolar lumen. Whereas human lysosomes are specialised for degradation, vacuoles have more diverse function including storage, sequestration and osmoregulation (Alberts, 2002).

Chapter 4: Expression Profile and Sub-cellular Localisation

4.1 Introduction

In this chapter I will describe the tissue-specific and sub-cellular localisation of the *AtPQL* family. Such information is crucial as it will give an insight into where *AtPQL* proteins function and is important for assessing potential roles that these proteins may fulfil.

In addition, I will present data concerning the transcriptional regulation of family members in response to external stimuli. Again this information may prove useful in placing the *AtPQL* family within specific transcriptional response pathways and pinpointing conditions under which these proteins may be important.

4.1.1 Expression profile

In this thesis, *AtPQL* transcript levels were investigated using a reverse-transcription real-time quantitative PCR (qPCR) approach.

To investigate tissue specific-expression, four different tissue types (roots, rosette leaves, floral buds and open flowers) of mature wild-type *A. thaliana* plants were examined. Particular attention was paid to flowers because microarray databases show higher expression of some family members in floral tissues (see section 4.2.2).

Tissue-specific localisation of one family member, *AtPQL1*, was further examined using stable transformants expressing the GUS reporter gene under the control of the *AtPQL1* promoter (see figure 4-1).

In addition, transcript level for various family members was checked at regular intervals over the course of a day and in response to two external stimuli - viral infection and wounding.

For viral infection and wounding experiments, the possible involvement of COI1 in the transcriptional response of *AtPQL* genes was investigated using *coil-16* mutant plants. *COI1* encodes a ubiquitin ligase complex (SCF^{COI1}) subunit involved in jasmonic acid (JA) signalling (Xie et al., 1998; Xu et al., 2002). It is thought that SCF^{COI1} induces

the expression of JA responsive genes by degrading negative regulators at the onset of JA signalling. As a result, *coi1-16*, shows greatly reduced sensitivity to JA (Ellis and Turner, 2002).

4.1.2 Sub-cellular localisation

To determine the sub-cellular localisation of *AtPQL1-6* a confocal microscopy approach was used. The coding region of each gene was fused to GFP (see figure 4-1), transiently expressed in tobacco epidermal cells by agroinfiltration and examined for fluorescence.

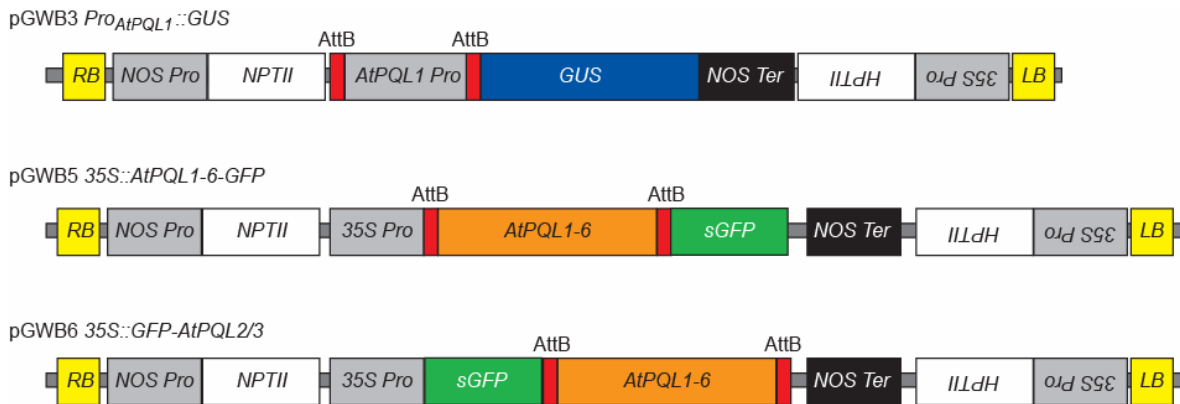


Figure 4-1. Expression cassettes of plasmids used in this chapter. The diagram shows the T-DNA cassette of the pGWB3, pGWB5 and pGWB6 vectors. In each case the Gateway attB recombination sites and plant selectable markers (HPTII and NPTII confer resistance to hygromycin and kanamycin respectively) are indicated. For pGWB3 *Pro_{AtPQL1}::GUS*, the position of the *AtPQL1* promoter relative to the *GUS* reporter gene and NOS terminator is shown. For pGWB5 and pGWB6 vectors, the position of the *AtPQL* coding region relative to the sGFP tag, CaMV 35S promoter and NOS terminator is shown.

4.2 Results

4.2.1 Expression profile

I will now describe the individual expression pattern of each gene according to the sub-groups defined in chapter three.

4.2.1.1 Sub-group 1: *AtPQL1*, *AtPQL2* and *AtPQL3*

Figure 4-2 shows the expression level of *AtPQL1-3* in four different tissues of wild-type *A. thaliana* plants relative to a stably expressed reference gene (*YLS8*) as determined by qPCR. *AtPQL1* and *AtPQL2* had the lowest overall expression level of all family members and were predominantly expressed in floral tissues. *AtPQL1* had the most specific pattern of expression, being found almost exclusively in floral buds. Very little transcript was found elsewhere (approximately 18 fold less in open flowers) although some expression was detectable (negative controls gave no signal). In contrast *AtPQL2* was most strongly expressed in open flowers but also showed relatively strong expression in roots and floral buds.

To further investigate the expression of *AtPQL1*, a 1 kb region immediately 5' of the start codon was fused to the GUS reporter (see figure 4-1) and used to make transgenic plants. As shown in figure 4-3, two independent T₂ lines, *Pro_{AtPQL1}::GUS-1* (figure 4-3 a-d) and *Pro_{AtPQL1}::GUS-2* (figure 4-3 e-h), showed a very similar pattern of GUS staining. GUS staining was consistently seen in pollen. This was observed in anthers containing pollen (figure 4-3b, e, f, h), pollen adhered to stigmas (figure 4-3c, g), and isolated pollen (figure 4-3a). GUS staining was not found in any other parts of the flower.

In sporophytic tissues, GUS expression was rarely observed. Despite checking several plants, no GUS staining was found in rosette leaves, stems, cauline leaves or siliques. Occasionally, weak GUS staining was seen in roots, close to the root shoot apex (see figure 4-3d), but this was only seen in two out of nine plants screened for line *Pro_{AtPQL1}::GUS-1* and was not observed at all in line *Pro_{AtPQL1}::GUS-2*.

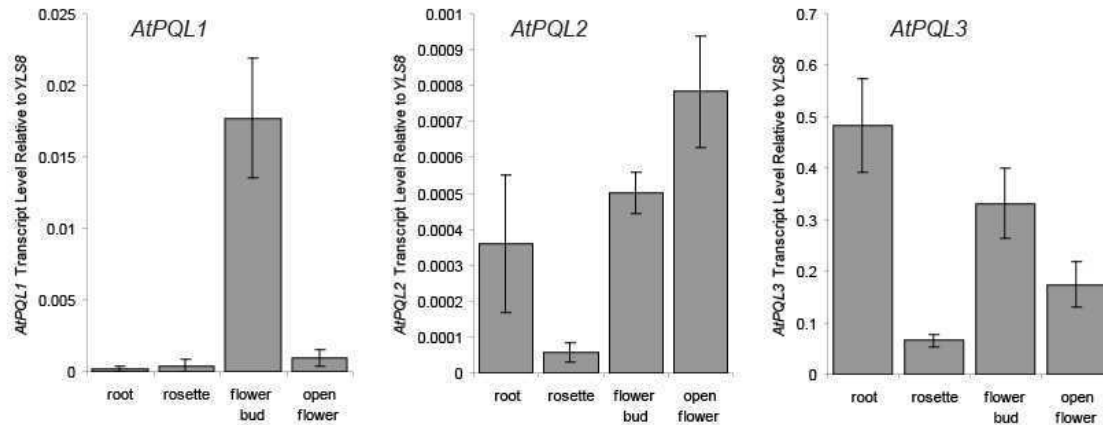
The tissue-specific expression pattern of the other member of this subgroup, *AtPQL3*, is shown in figure 4-2a. *AtPQL3* differed from *AtPQL1* and *AtPQL2*, in that it was not preferentially expressed in flowers. In addition, the overall transcript level of *AtPQL3* was much higher. Peak transcript level was found in roots with lower levels in rosette leaves and flowers.

Figure 4-2b and c shows the transcriptional response of *AtPQL3* to viral infection and wounding in the rosette leaves of mature hydroponically grown plants. *AtPQL1* and *AtPQL2* were omitted from this experiment due to their low level of expression in rosette leaves. The response to viral infection was assessed by inoculating plants with cauliflower mosaic virus (CaMV) by particle bombardment and then harvesting infected leaves six days later. The response to wounding was checked by piercing each leaf in a rosette two to three times with a fine needle and harvesting three hours later. In both cases mock treated samples were used as controls (see materials and methods). The transcriptional response of *AtPQL3* to both treatments was measured in both wild-type and *coil-16* mutant plants.

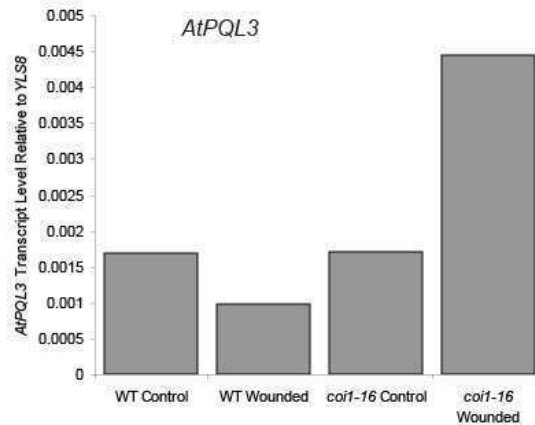
Expression of *AtPQL3* was not affected by viral infection of wild-type plants. However, in the *coil* background a 3.3-fold increase in transcript level was observed. Similarly in response to wounding no induction, and even a slight repression, was seen in wild type, but a 2.6-fold increase in *AtPQL3* transcript level was seen in *coil-16*.

No reproducible diurnal regulation of *AtPQL3* expression could be seen in long or short day conditions (data not shown).

A. Tissues



B. Wounding



C. Viral Infection

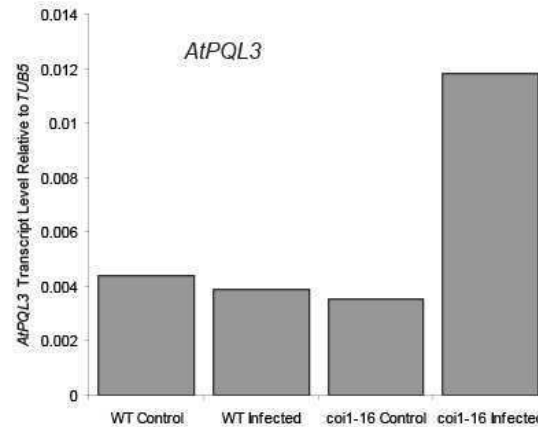


Figure 4-2. Expression level of *AtPQL1-3* as determined by real-time reverse-transcription PCR. In each case, RNA was extracted and used to synthesize cDNA which was used as a template for PCR. In **A** roots, rosette leaves, floral buds, or open flowers were harvested from six-week old plants. In **B** hydroponically grown wild type or *coi1-16* mutant plants were harvested three hours after being wounded by piercing leaves with a hypodermic needle. Controls were treated in the same way but not wounded. In **C** wild type or *coi1-16* mutant plants were harvested six days after infection with cauliflower mosaic virus via particle bombardment. Controls were subject to particle bombardment with virus-free particles. *YLS8* or *TUB5* was used as a stably expressed reference gene as indicated. In each case, the relative transcript level is the ratio of the absolute amounts of the indicated *AtPQL* gene to the reference gene in each cDNA sample as determined by comparison to a standard curve produced from samples of known template concentration. The values given in **A** are the average of 3 biological replicates (each biological replicate is the average of two technical replicates). Error bars represent standard error. The values given in **B** and **C** represent a single biological replicate.

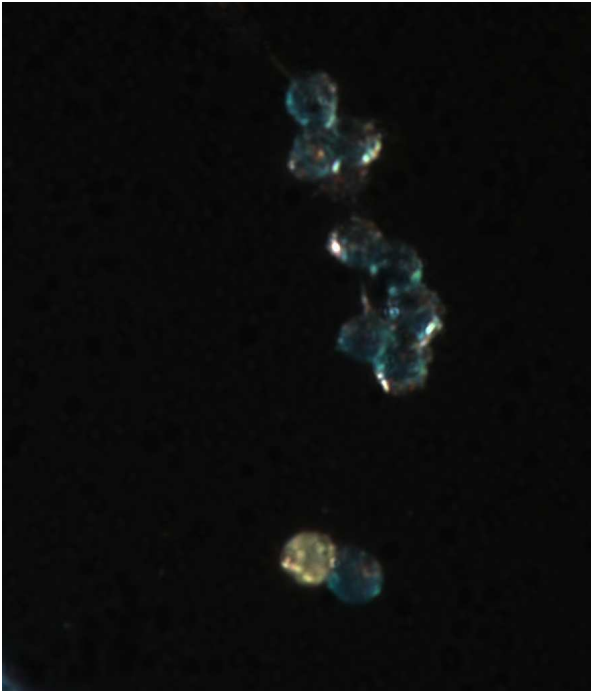
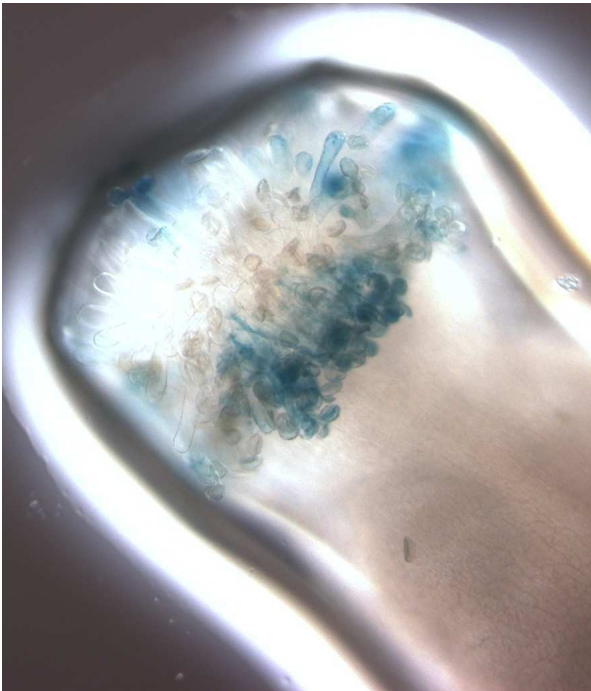
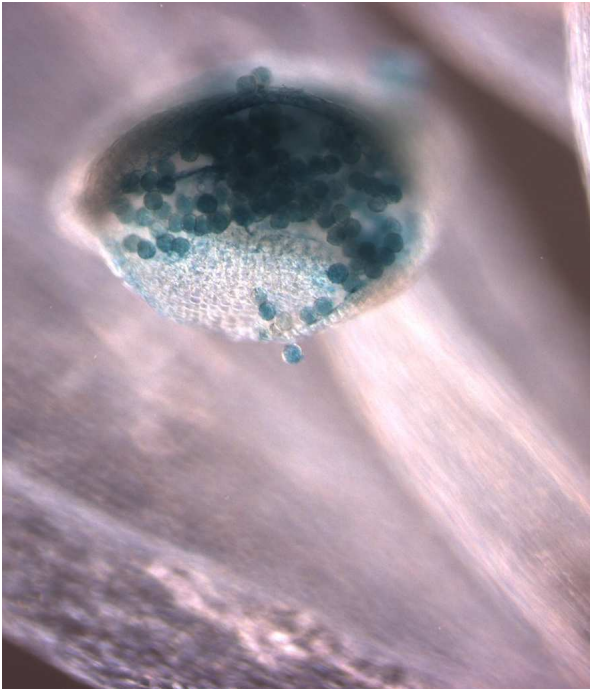
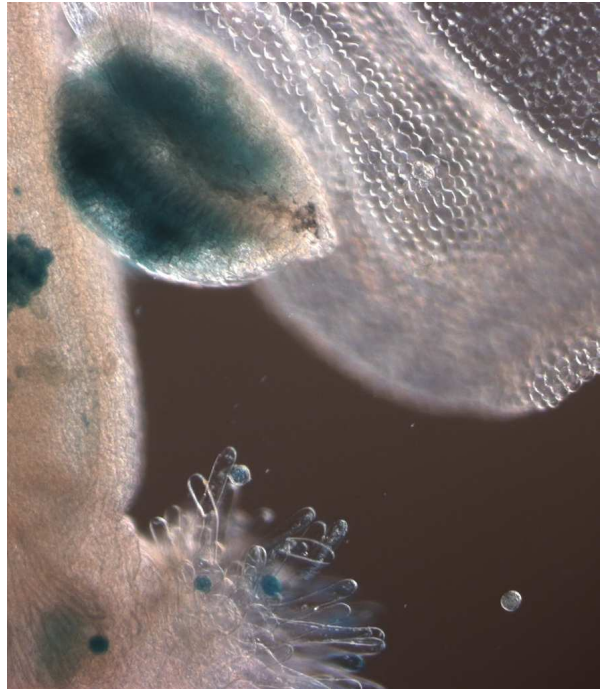
A**B****C****D**

Figure 4-3. Histochemical GUS staining of transgenic plants expressing an *AtPQL1* promoter::GUS construct. Images are of two independent T₂ transgenic lines; *Pro_{AtPQL1}::GUS-1* (A-D) and *Pro_{AtPQL1}::GUS-2* (E-H). Tissue was stained for approximately 16 hours in 2mM X-Gluc and chlorophyll removed by washing in ethanol. GUS staining was visible in isolated pollen (A), pollen filled anthers (B,C,E,F,H), pollen attached to stigmas (C,G), and roots (D)

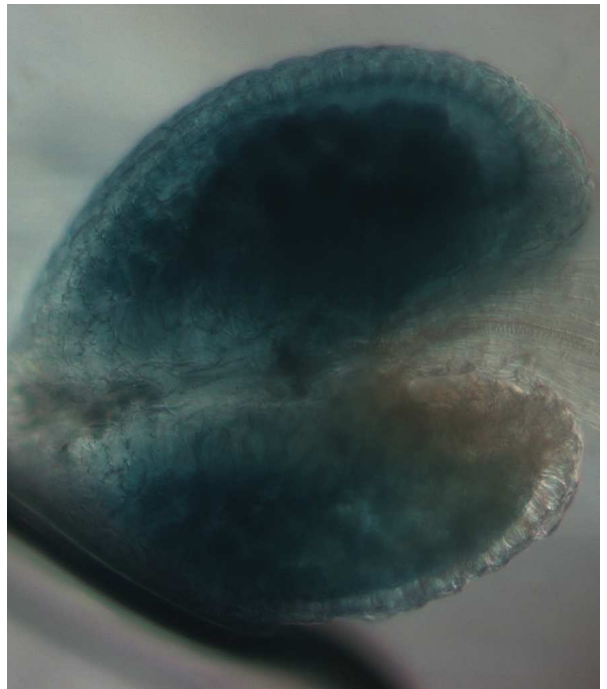
E



F



G



H

Figure 4-3. Continued from previous page.

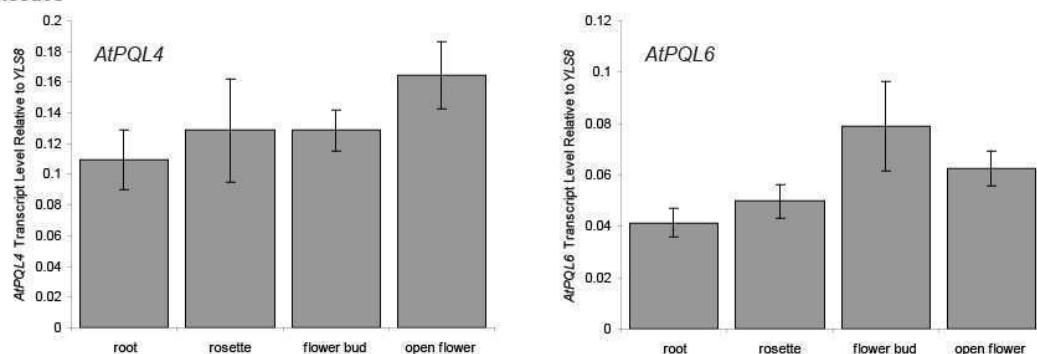
4.2.1.2 Sub-group 2: *AtPQL4* and *AtPQL6*

The expression level of *AtPQL4* and *AtPQL6* in different tissues is shown in figure 4-4a. The relative transcript levels of the two genes were similar. Transcript level relative to *YLS8* ranged from 0.1 to 0.16 for *AtPQL4* and from 0.04 to 0.08 for *AtPQL6*. Both genes showed little variation in expression level between different tissues. *AtPQL4* had slightly higher transcript levels in open flowers and slightly lower levels in roots. *AtPQL6* showed highest expression in floral buds.

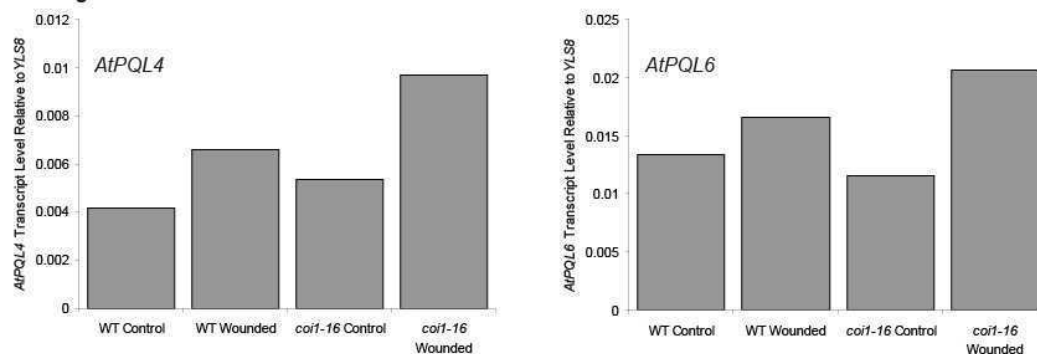
The transcript levels of both *AtPQL4* and *AtPQL6* were increased upon viral infection and wounding (figure 4-4b and c). Unlike *AtPQL3*, which was induced by viral infection and wounding only in the *coil-16* mutant background, both *AtPQL4* and *AtPQL6* were induced in both *coil-16* and wild-type. However, the response to both stresses was enhanced in *coil-16*. *AtPQL4* transcript level was increased 1.6-fold by viral infection in wild type and 2.2-fold in *coil-16*. *AtPQL6* transcript level was increased 1.8-fold in wild type and 2.7-fold in *coil-16*. Likewise in response to wounding, *AtPQL4* transcript level was increased 1.6-fold in wild type and 1.8-fold in *coil-16*. *AtPQL6* transcript level was increased 1.2-fold in wild type and 1.8-fold in *coil-16*.

To investigate diurnal regulation, seedlings were grown on agar plates in short or long-day conditions for 12 days before being harvested at regular intervals over the course of a day. As shown in figure 4-4d and e, both *AtPQL4* and *AtPQL6* show diurnal changes in expression level. In both short and long-day photoperiods, both genes show increasing expression during the light period. In long-day conditions a steady increase can be seen from the morning just after the lights were switched on until the evening. *AtPQL4* showed peak expression at midnight just after the lights were switched off whereas *AtPQL6* showed peak expression at 8 p.m. and a slight decrease in expression by midnight. In short days, a similar pattern was seen, with *AtPQL4* showing peak expression at 8 p.m., the first time point after the lights were switched off, and *AtPQL6* showing peak expression at 4 p.m., the last time point whilst the lights were on.

A. Tissues



B. Wounding



C. Viral Infection

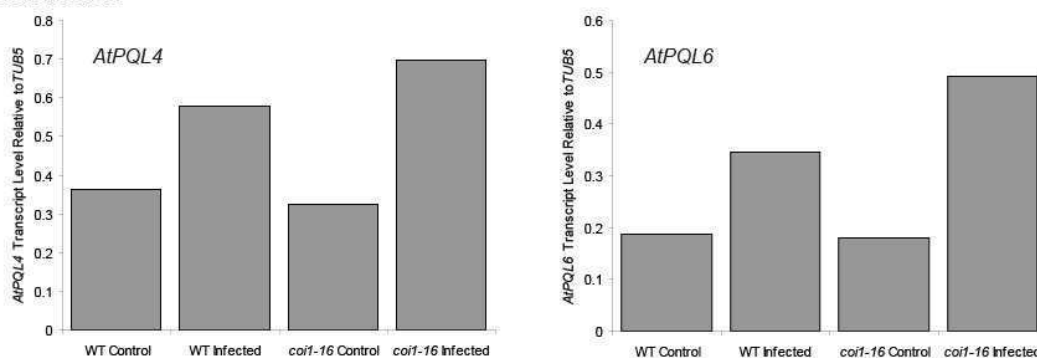
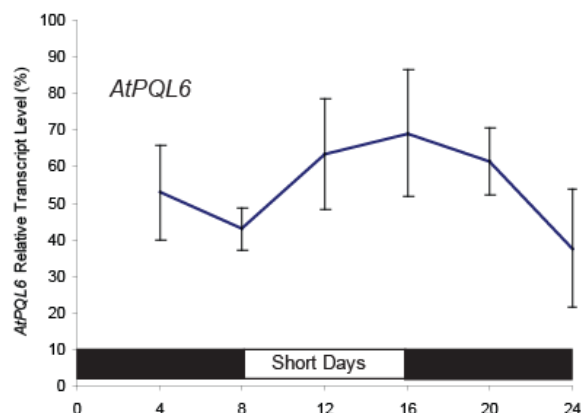
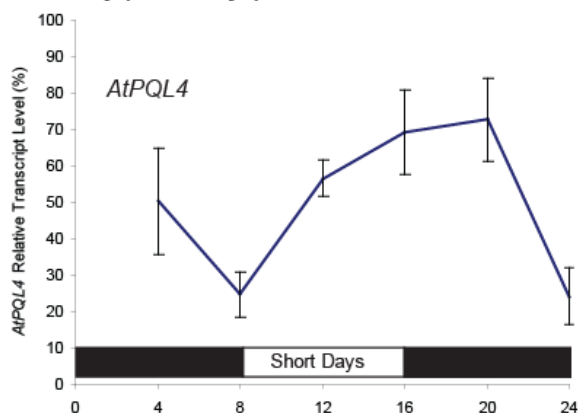


Figure 4-4. Transcript level of *AtPQL4* and *AtPQL6* as determined by real-time reverse-transcription PCR. In each case, RNA was extracted and used to synthesize cDNA which was used as a template for PCR. In **A** roots, rosette leaves, floral buds, or open flowers were harvested from six-week old plants. In **B** hydroponically grown wild type or *coi1-16* mutant plants were harvested three hours after being wounded by piercing leaves with a hypodermic needle. Controls were treated in the same way but not wounded. In **C** wild type or *coi1-16* mutant plants were harvested six days after infection with cauliflower mosaic virus via particle bombardment. Controls were subject to particle bombardment with virus-free particles. In **D** and **E** 12-day old seedlings on agar plates were harvested at four hour intervals over the course of a day in short-day (**D**) or long-day (**E**) photoperiods. *YLS8* or *TUB5* was used as a stably expressed reference gene as indicated. White and black bars indicate light and dark periods in long and short-day conditions. **Continued overleaf.**

D. Time of Day (Short Days)



E. Time of Day (Long Days)

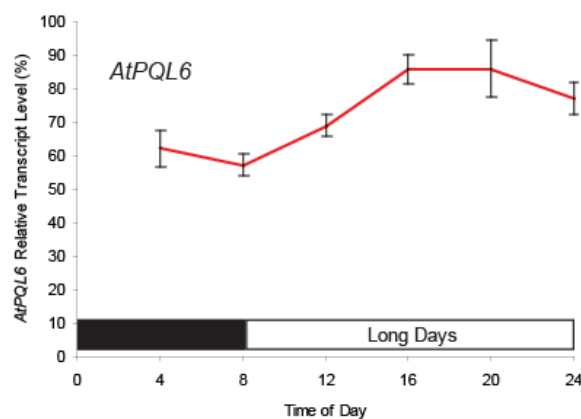
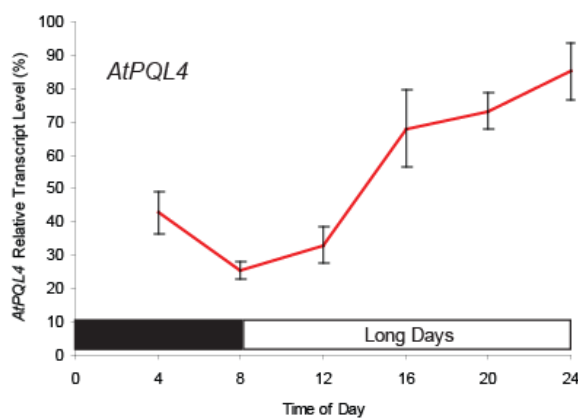


Figure 4-4. Continued from previous page. In each case, the relative transcript level is the ratio of the absolute amounts of the indicated *AtPQL* gene to the reference gene in each cDNA sample as determined by comparison to a standard curve produced from samples of known template concentration. The values given in **A**, **D** and **E** are the average of 3 biological replicates (each biological replicate is the average of two technical replicates). Error bars represent standard error. The values given in **B** and **C** represent a single biological replicate.

4.2.1.3 Sub-group 3: *AtPQL5*

The expression level of *AtPQL5*, the sole member of sub-group 3, is shown in figure 4-5. In mature plants, *AtPQL5* was found to be the most strongly expressed family member overall. Transcript level was highest in open flowers but expression was also relatively high elsewhere in the plant. Next highest transcript level was found in rosette leaves with lower levels in roots and floral buds.

Like other *AtPQL* genes, *AtPQL5* expression was induced by both viral infection and wounding. In wild-type plants, transcript level was increased 1.7-fold by viral infection and 2-fold by wounding. Like *AtPQL4* and *AtPQL6*, the response to both treatments was enhanced in the *coil-16* background. In *coil-16*, transcript level was increased 2-fold by viral infection and 2.2-fold by wounding.

There were also changes in the expression of *AtPQL5* over the course of a day. Similar to *AtPQL4* and *AtPQL6*, expression of *AtPQL5* was found to be higher during the day than the night. In short days, a clear diurnal pattern was seen with expression highest at noon and lowest at midnight. In long days, expression increased over the course of the day appearing to peak at 4 p.m. and then dipping at 8 p.m. before increasing again at midnight, just after the lights were switched off.

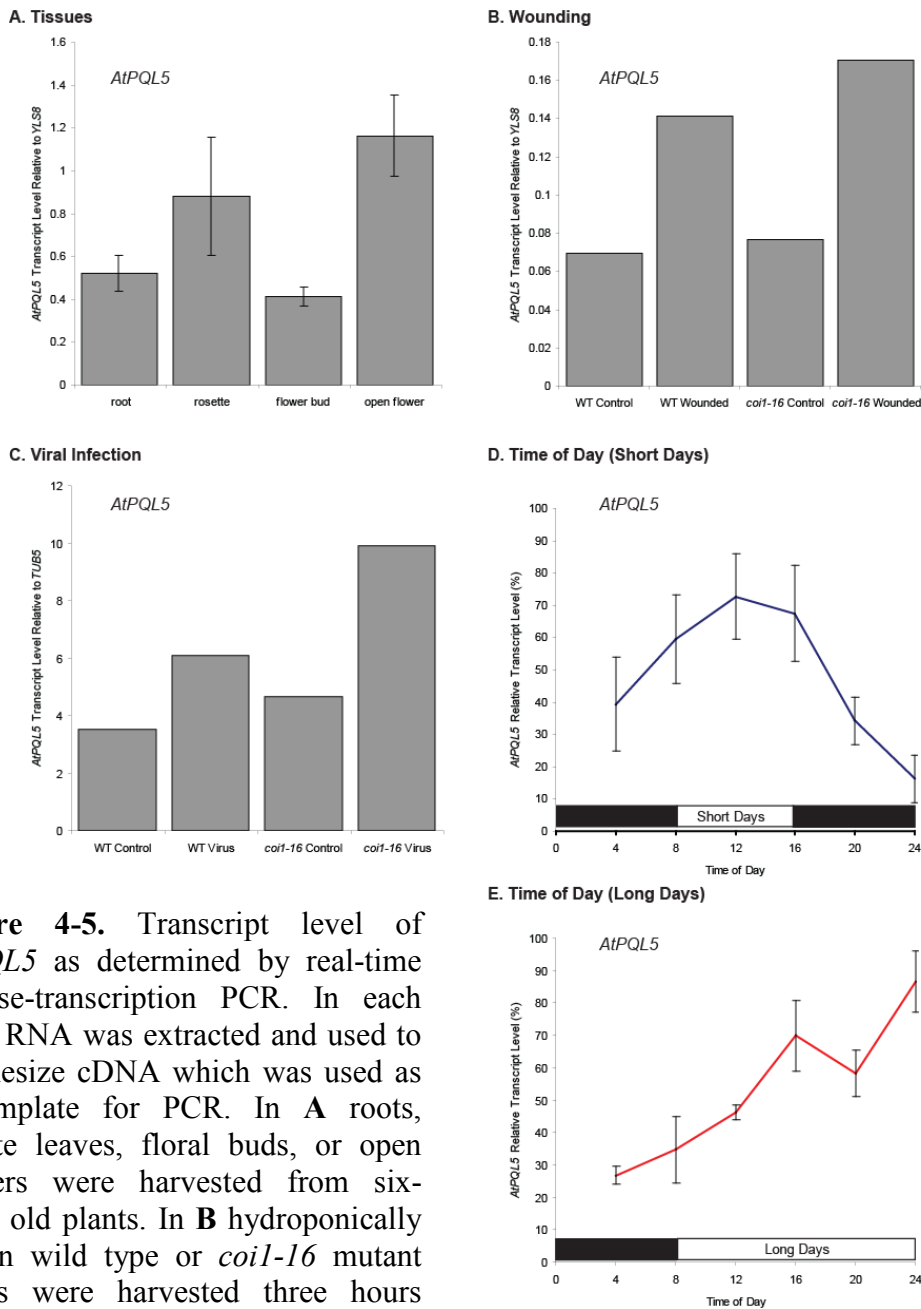


Figure 4-5. Transcript level of *AtPQL5* as determined by real-time reverse-transcription PCR. In each case, RNA was extracted and used to synthesize cDNA which was used as a template for PCR. In **A** roots, rosette leaves, floral buds, or open flowers were harvested from six-week old plants. In **B** hydroponically grown wild type or *coi1-16* mutant plants were harvested three hours after being wounded by piercing leaves with a hypodermic needle. Controls were treated in the same way but not wounded. In **C** wild type or *coi1-16* mutant plants were harvested six days after infection with cauliflower mosaic virus via particle bombardment. Controls were subject to particle bombardment with virus-free particles. In **D** and **E** 12-day old seedlings on agar plates were harvested at four hour intervals over the course of a day in short-day (**D**) or long-day (**E**) photoperiods. *YLS8* or *TUB5* was used as a stably expressed reference gene. White and black bars indicate light and dark periods in long and short-day conditions. In each case, the relative transcript level is the ratio of the absolute amounts of *AtPQL5* to the reference gene in each cDNA sample as determined by comparison to a standard curve produced from samples of known template concentration. The values given in **A**, **D** and **E** are the average of 3 biological replicates (each biological replicate is the average of two technical replicates). Error bars represent standard error. The values given in **B** and **C** represent a single biological replicate.

4.2.2 Microarray databases

This section will summarise published microarray data regarding the expression of the *AtPQL* family. Unfortunately, neither *AtPQL3* nor *AtPQL4* are represented on the Affymetrix ATH1 22K probeset so expression data is limited. For the other family members data was obtained from Genevestigator (Zimmermann et al., 2004), the *DIURNAL* microarray viewer (Mockler, 2007) and eFP browser (Winter et al., 2007).

4.2.2.1 Microarray data - tissue-specific localisation

Microarray data for tissue-specific expression was obtained from Genevestigator (Zimmermann et al., 2004) and is shown in figure 4-6. This data and the qPCR experiments presented here show a similar expression profile. For *AtPQL1*, *AtPQL2* and *AtPQL5* expression was highest in pollen and stamen (stamen samples include pollen in non-dehiscent anthers). Signal intensity values for *AtPQL1* were up to twelve-fold higher in pollen (623) and stamen (862) than in vegetative tissues (average signal intensity excluding pollen and stamen was 143 indicating little or no expression). *AtPQL2* also gave highest signal intensity values for stamen (1124) and pollen (1165) although some vegetative tissues also gave relatively high values including root endodermis (609) and root stele (444). The average signal intensity for *AtPQL5* (4826) was much stronger than for *AtPQL1* (173) and *AtPQL2* (282). Like *AtPQL1* and *AtPQL2*, highest values were from pollen (60324) and stamen (15968) with relatively even expression values elsewhere (average signal intensity excluding pollen and stamen was 3072).

Again in agreement with the qPCR results presented here, microarray data shows expression of *AtPQL6* across a wide range of tissues (average signal intensity of 2090). Signal intensity values were lowest for ovaries (937) and lateral root caps (989) with highest values for lateral root tips (3284), senescent leaves (3275) and cell suspension cultures (3337).

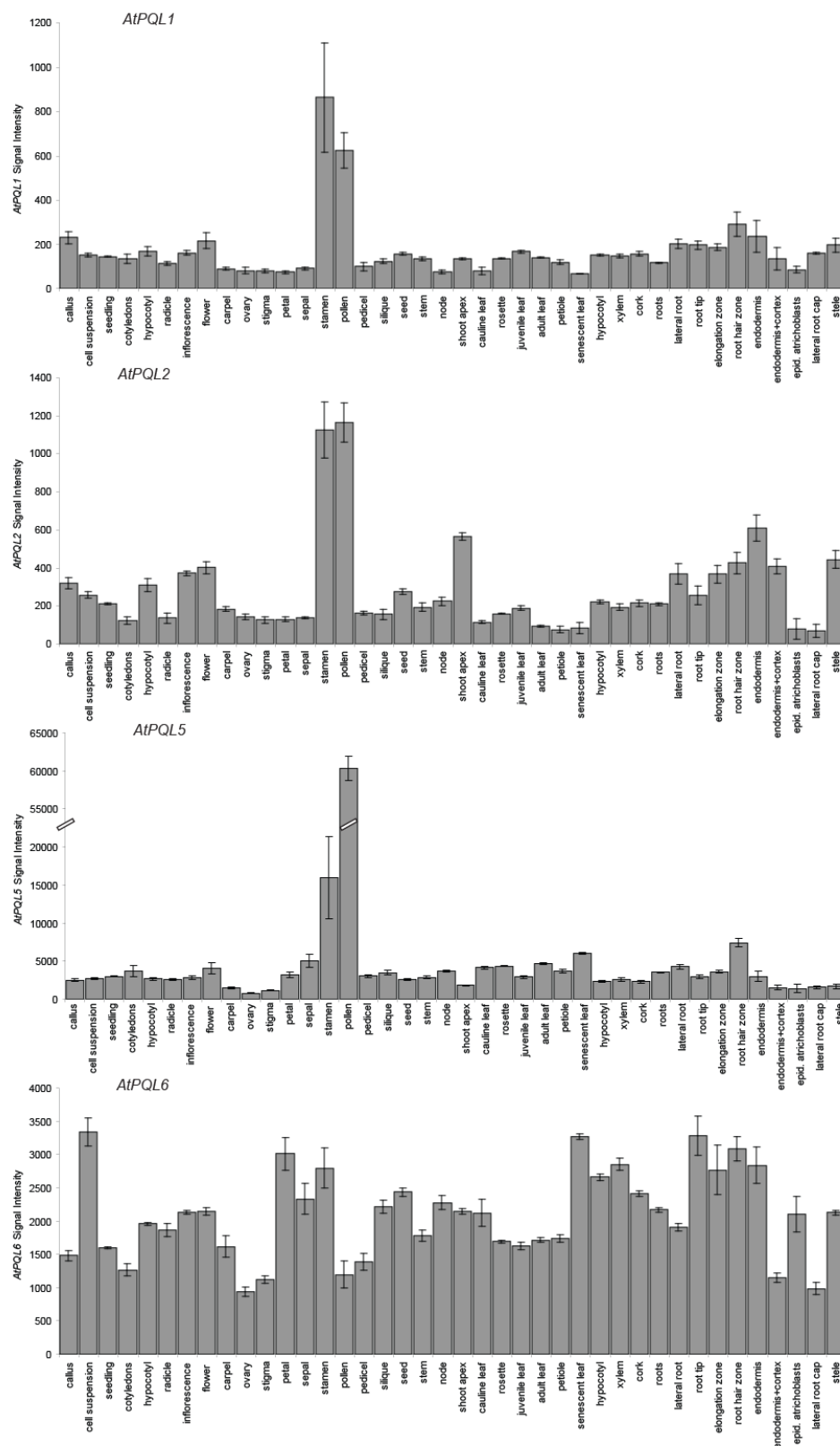


Figure 4-6. Microarray expression data for *AtPQL1*, *AtPQL2*, *AtPQL5* and *AtPQL6* in different tissues. Data was obtained from microarray experiments collated at the Genevestigator public microarray database (Zimmermann et al., 2004). All experiments used the affymetrix ATH1 with RNA extracted from the indicated tissue. Data is reported as absolute signal intensity values. The number of biological replicates vary from 2 (pollen) to 828 (adult leaf). Error bars represent standard error.

4.2.2.2 Microarray data - expression during pollen development

As three of the four *AtPQL* genes represented on the ATH1 array showed highest transcript levels in pollen or stamen, expression during pollen development was examined using eFP browser (Winter et al., 2007). Figure 4-7 shows data from pollen divided into four groups based on developmental stage. These groups were uninucleate microspores, binucleate pollen, tricellular pollen and mature pollen grains.

Expression varied amongst the different stages of development for all four genes. *AtPQL1* and *AtPQL6* both showed highest signal intensity values in uninucleate microspores and lower values in the later stages of pollen development. The ratio of signal intensity in uninucleate microspores compared to mature pollen was 3.6 for *AtPQL1* and 5.7 for *AtPQL6*.

In direct contrast, *AtPQL2* and *AtPQL5* showed peak expression in mature pollen grains. *AtPQL2* showed the least variation in signal intensity, but nevertheless a gradual increase in expression is seen during development from bicellular to mature pollen. *AtPQL5* shows the most pronounced changes in expression during pollen development, with signal intensity in mature pollen 25.6-fold higher than in uninucleate microspores.

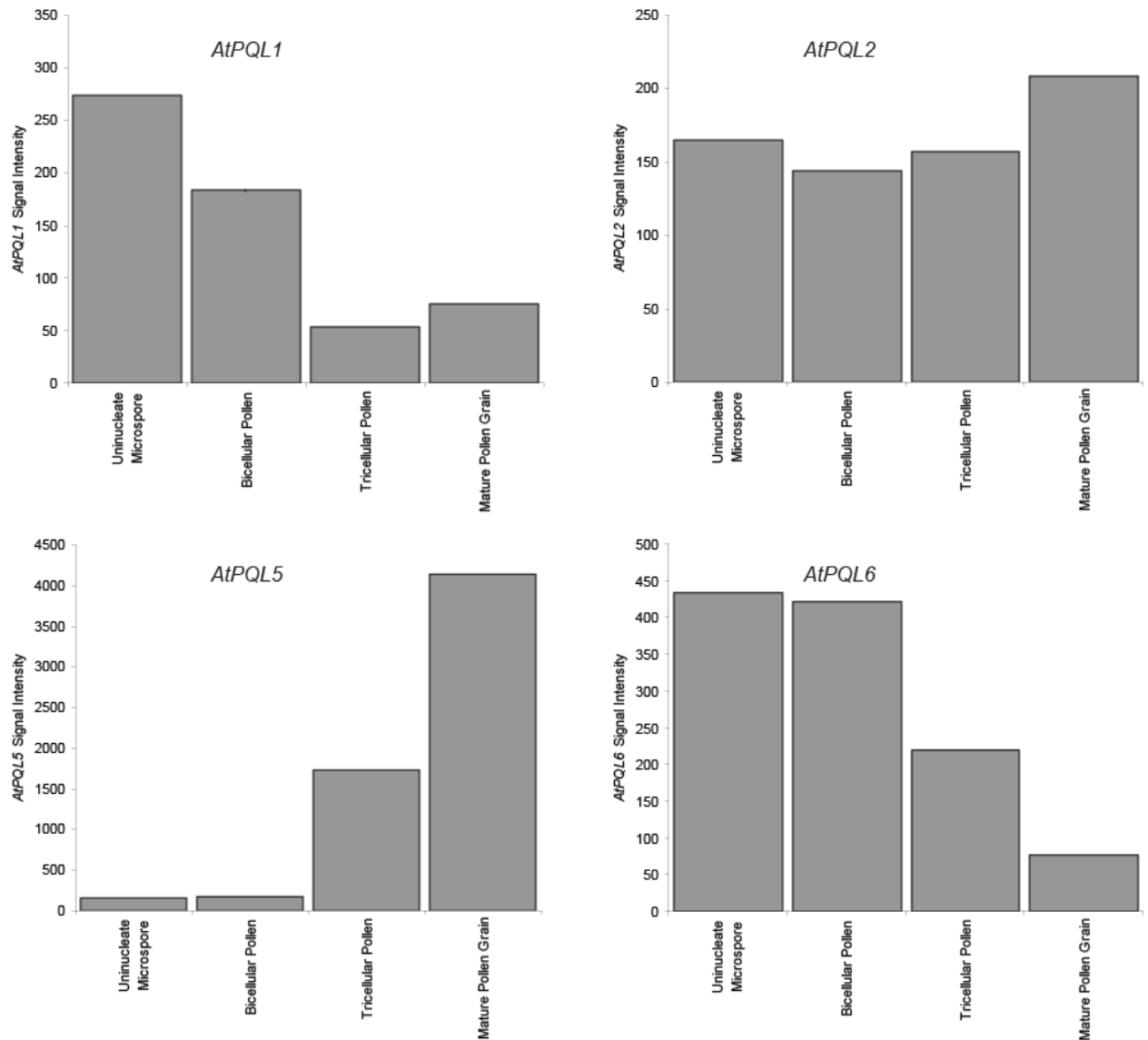


Figure 4-7. Microarray expression data for *AtPQL1*, *AtPQL2*, *AtPQL5* and *AtPQL6* in different tissues. Data was obtained from microarray experiments collated at eFP browser (Winter et al., 2007). The experiments presented used RNA extracted from *Ler0* pollen sorted into four different developmental stages (Honys and Twell, 2004). In each case the affymetrix ATH1 array was used. Data is reported as absolute signal intensity values. Data for uninucleate microspores, bicellular pollen and tricellular pollen is an average of two biological replicates. Data for mature pollen grains is from a single biological replicate.

4.2.2.3 Microarray data - diurnal changes in expression

Changes in expression level over the course of a day were investigated using the *DIURNAL* microarray viewer (Mockler, 2007). Data for *AtPQL5* and *AtPQL6* are shown in figure 4-8 (*AtPQL1* and *AtPQL2* were not included due to their low level of expression in vegetative tissue). A diurnal rhythm in expression is apparent for both genes and similar to that shown here by qPCR. In long-day conditions, *AtPQL5* showed peak expression early in the day, decreasing expression during the afternoon and a slight recovery in late evening. In short-day conditions, expression peaked in the middle of the day and sharply decreased during the night. *AtPQL6* showed a very similar pattern in long and short days. Several spikes in transcript level were seen but in both photoperiods expression was higher during the day than during the night.

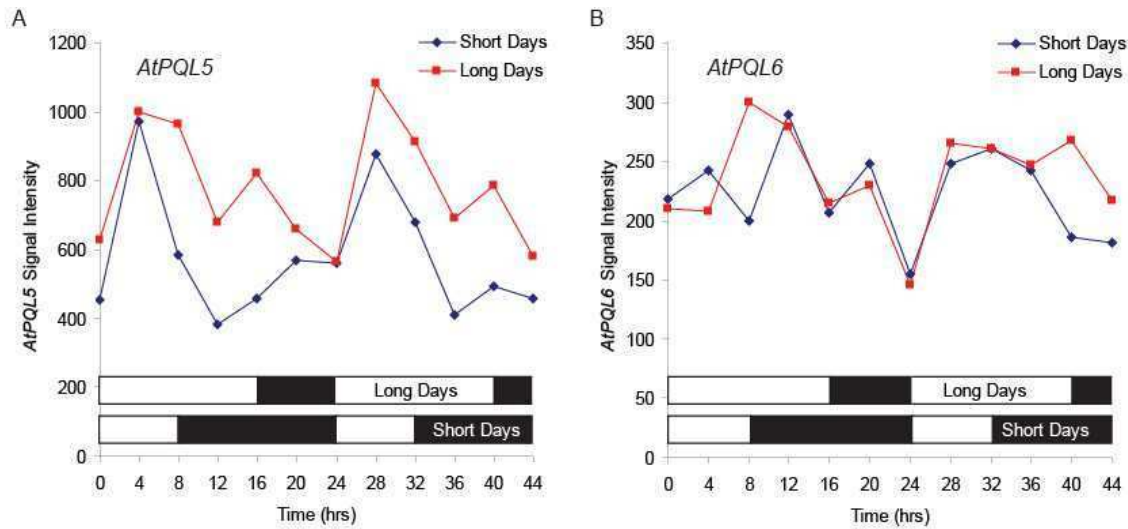


Figure 4-8. Microarray expression data for *AtPQL5* (A) and *AtPQL6* (B) over the course of two days. Data was obtained from microarray experiments collated at the *DIURNAL* database (Mockler, 2007). The experiments presented used RNA extracted from *Ler0* seedlings grown on agar plates with Murashige and Skoog medium and 3% sucrose. Seedlings were grown in either long-day (16hr light/ 8hr dark) or short day (8hr light/16hr dark) conditions for 7 days before harvesting commenced. Plants were harvested at four hour intervals over the course of 48 hours. Data is reported as absolute signal intensity values and represents a single replicate for each time point.

4.2.3 Sub-cellular localisation

I will now describe the sub-cellular localisation of each protein as determined by confocal microscopy following transient expression of GFP fusion constructs in *Nicotiana tabacum*. The T-DNA cassettes of the expression vectors used are shown in figure 4-1.

4.2.3.1 Sub-group 1: *AtPQL1*, *AtPQL2* and *AtPQL3*

Images of leaf epidermal cells expressing *AtPQL1* fused to GFP at the C-terminal can be seen in figure 4-9. GFP fluorescence was seen in the majority of infiltrated cells. In most cases, fluorescence was seen close to the edge of each cell. As indicated by the white arrows in the right-hand panels of figure 4-9, rather than being at the very edge, fluorescence was often seen in a smooth line distinct from the cell periphery. This suggests localisation to the tonoplast. Often, the distinction was particularly clear when a chloroplast (visible due to autofluorescence) was present between the vacuole and plasma membrane (figure 4-9, top-right panel). Such images are similar to those previously shown for tonoplast localised proteins (Cutler et al., 2000; Loque et al., 2005).

Plasmids were also constructed to express *AtPQL2* and *AtPQL3*-GFP fusions. However, GFP fluorescence could not be detected for either construct, regardless of whether the tag was placed at the N or C-terminal of the coding region. Attempts to optimise expression by changing the density of the infiltration culture were unsuccessful.

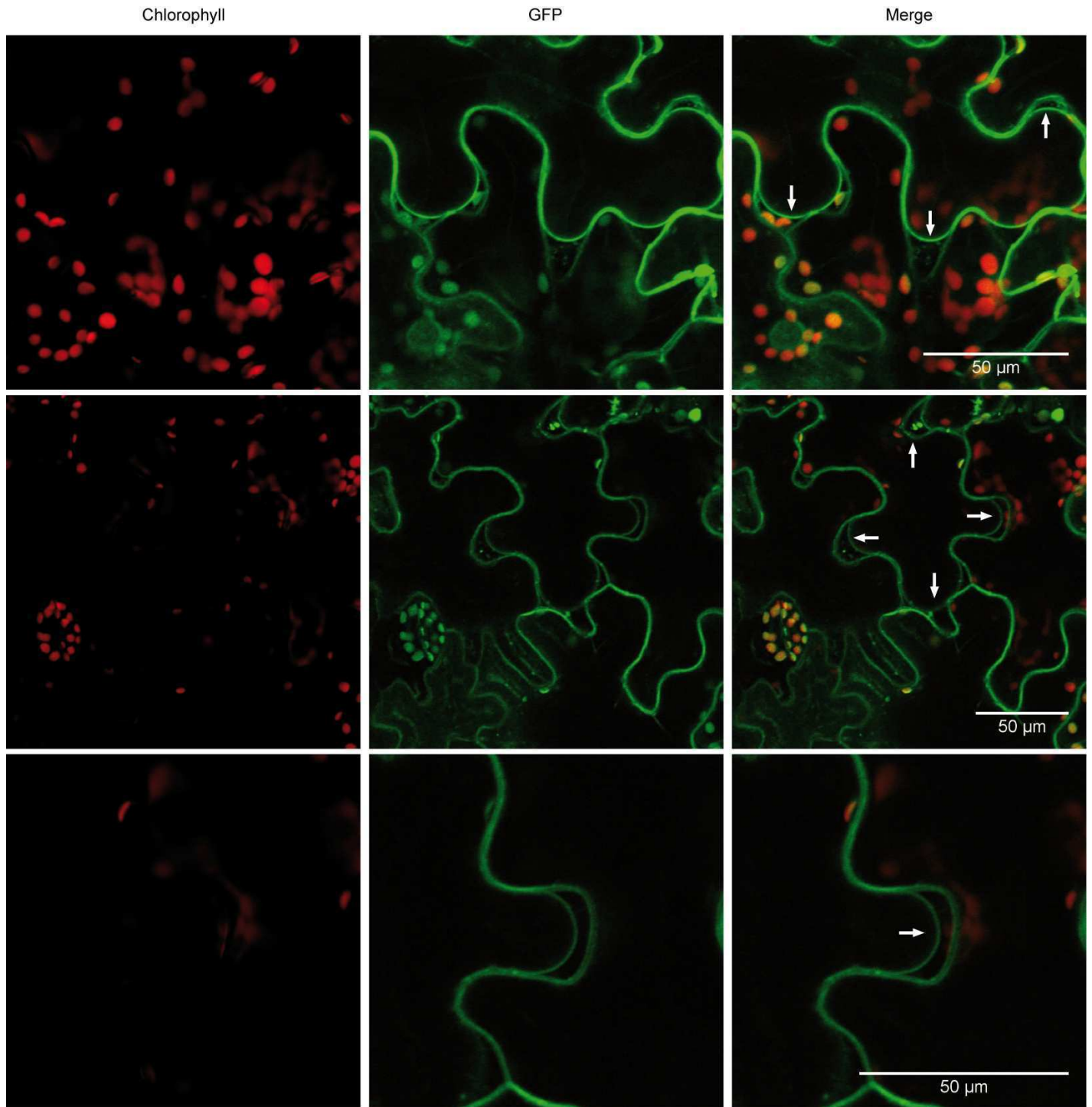


Figure 4-9. Sub-cellular localisation of *AtPQL1*. The coding region of *AtPQL1* was fused to GFP and transiently expressed in tobacco leaf epidermal cells by agroinfiltration. Tissue was examined two days later using confocal microscopy. Fluorescence was detected after excitation at 488 nm, using a 505 – 530 nm filter for GFP and a 560 – 615 nm filter for chloroplast autofluorescence. The panel on the right shows an overlay of the two images.

4.2.3.2 Sub-group 2: *AtPQL4* and *AtPQL6*

Images of tobacco epidermal cells transiently transformed with *AtPQL4* and *AtPQL6*-GFP fusion proteins are shown in figures 4-10 and 4-11. GFP fluorescence was obtained for both *AtPQL4* and *AtPQL6* but in contrast to *AtPQL1*, the majority of cells did not show fluorescence. For both constructs, transgene expression could be seen in approximately 25% of all the cells checked. Again unlike *AtPQL1*, neither construct showed a smooth line of fluorescence at the edge of cells. Instead the majority of fluorescence was seen in internal compartments forming net-like patterns often visible throughout the cell. The pattern of fluorescence is reminiscent of that previously observed for known ER localised proteins (Haseloff et al., 1997; Fluckiger et al., 2003).

4.2.3.3 Sub-group 3: *AtPQL5*

Tobacco epidermal cells expressing *AtPQL5*-GFP fusion constructs are shown in figure 4-12. Fluorescence was seen in almost all the cells that were examined. As with *AtPQL1* the majority of fluorescence was seen close to, but distinct from, the cell periphery (see white arrows) suggesting localisation to the tonoplast.

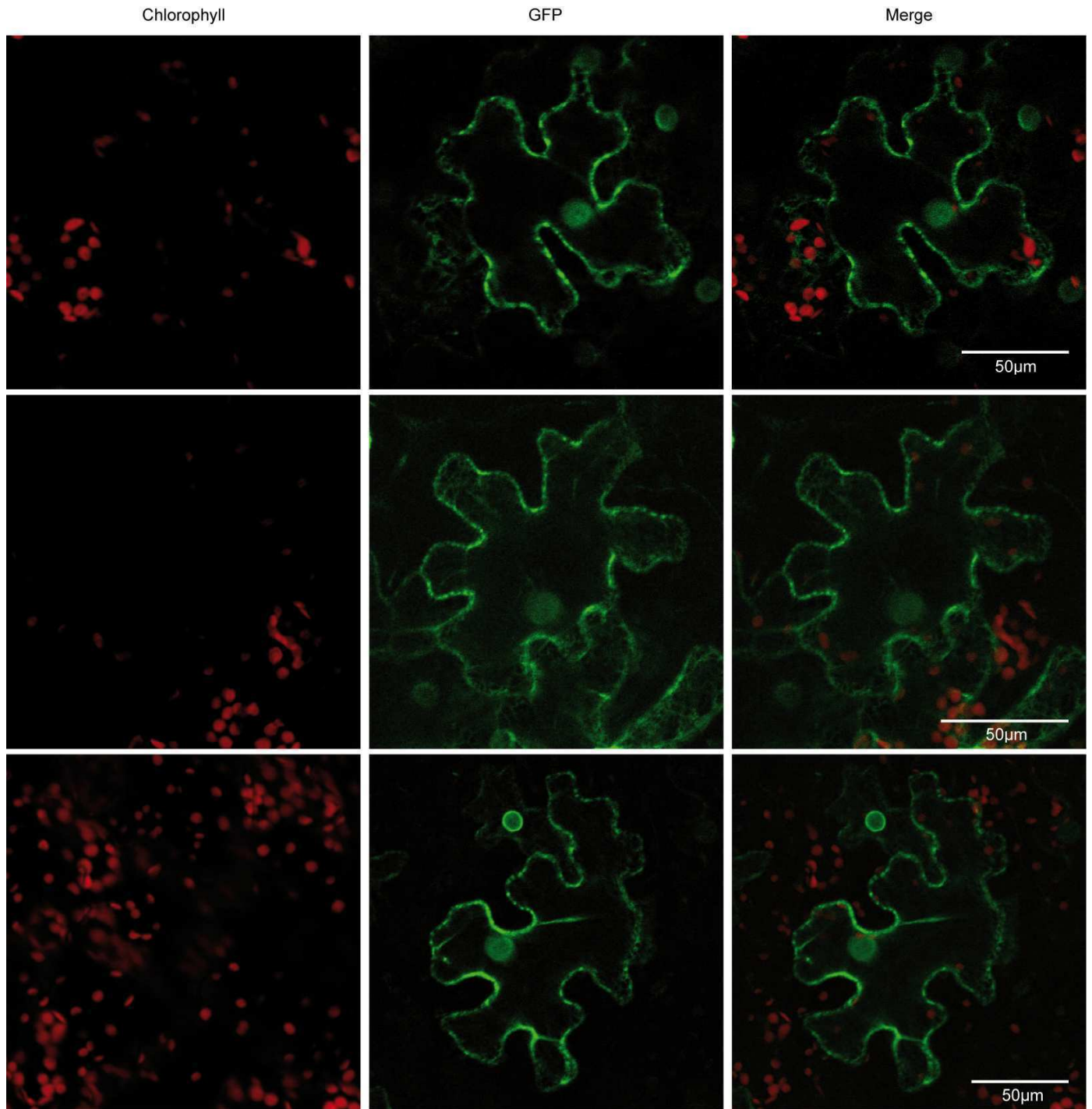


Figure 4-10. Sub-cellular localisation of *AtPQL4*. The coding region of *AtPQL4* was fused to GFP and transiently expressed in tobacco leaf epidermal cells by agroinfiltration. Tissue was examined two days later using confocal microscopy. Fluorescence was detected after excitation at 488 nm, using a 505 – 530 nm filter for GFP and a 560 – 615 nm filter for chloroplast autofluorescence. The panel on the right shows an overlay of the two images.

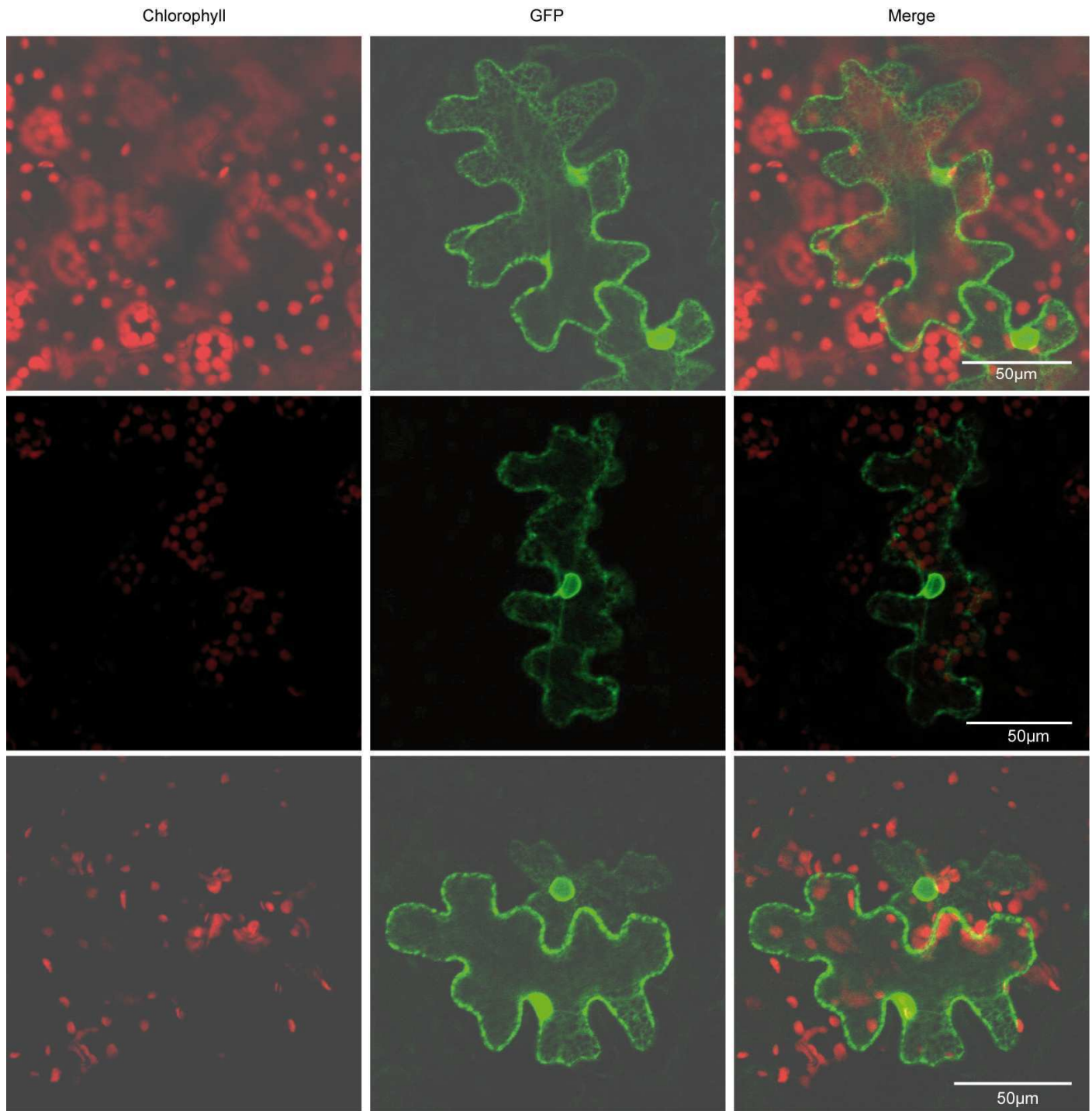


Figure 4-11. Sub-cellular localisation of *AtPQL6*. The coding region of *AtPQL6* was fused to GFP and transiently expressed in tobacco leaf epidermal cells by agroinfiltration. Tissue was examined two days later using confocal microscopy. Fluorescence was detected after excitation at 488 nm, using a 505 – 530 nm filter for GFP and a 560 – 615 nm filter for chloroplast autofluorescence. The panel on the right shows an overlay of the two images.

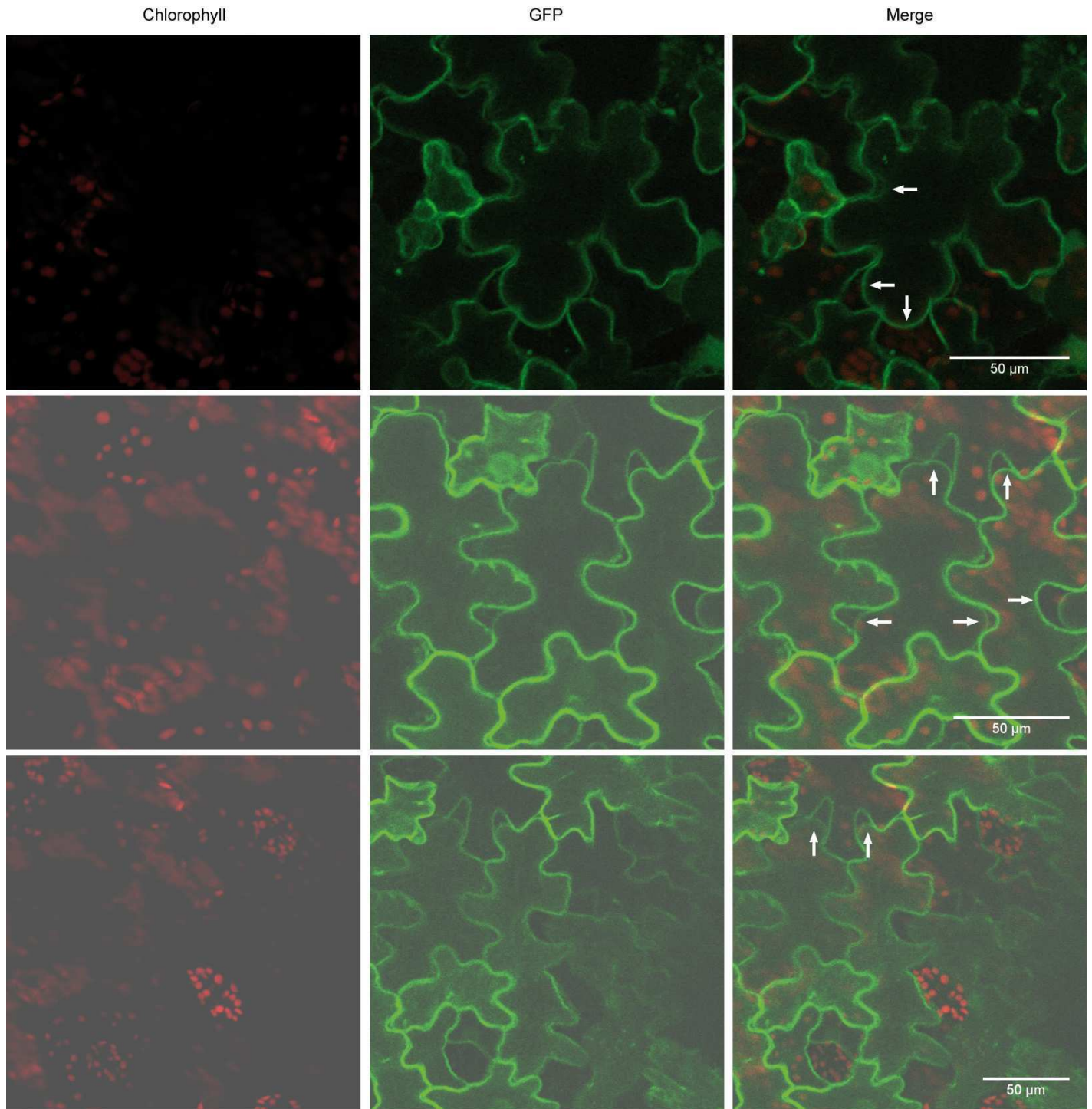


Figure 4-12. Sub-cellular localisation of *AtPQL5*. The coding region of *AtPQL5* was fused to GFP and transiently expressed in tobacco leaf epidermal cells by agroinfiltration. Tissue was examined two days later using confocal microscopy. Fluorescence was detected after excitation at 488 nm, using a 505 – 530 nm filter for GFP and a 560 – 615 nm filter for chloroplast autofluorescence. The panel on the right shows an overlay of the two images.

4.3 Discussion

This section will discuss the expression and localisation data presented here and its implications for the function of *AtPQL* proteins. I will first discuss the unique features of each family member and sub-grouping before discussing some of the common features shared by multiple family members.

4.3.1 Unique features of individual family members and sub-groups

4.3.1.1 Sub-group 1: *AtPQL1*, *AtPQL2* and *AtPQL3*

The expression patterns of *AtPQL1* and *AtPQL2* are similar. qPCR experiments presented here show that both are most strongly expressed in floral tissues. Promoter::GUS screening and microarray databases confirm this observation and show that the highest levels of expression are found in pollen. The fact that the two proteins show high levels of sequence homology (57% identity and 65% similarity at the protein level) and share similar regulation, suggests they may play a similar (possibly redundant) role. However, closer observation of microarray data suggests they are expressed at different stages of pollen development. In addition, whereas *AtPQL1* seems to be almost exclusively expressed in pollen, *AtPQL2* is less specific and has relatively high levels of expression elsewhere in the plant (though still very low in absolute terms).

The expression pattern of *AtPQL3* differs considerably from that of *AtPQL1* and *AtPQL2*. *AtPQL3* is expressed at a much higher level than *AtPQL1* and *AtPQL2*, and is not predominantly expressed in floral tissues. Instead, highest expression was found in root tissue.

Evidence from the expression of *AtPQL1*-GFP fusions suggests localisation to the tonoplast. In contrast, GFP fluorescence could not be detected for either *AtPQL2*-GFP or *AtPQL3*-GFP fusions regardless of whether the GFP tag was placed at the N- or C-terminus. It may be that further optimisation of culture density and infiltration buffer are required to encourage optimum T-DNA transfer and expression for these constructs. A common problem with transient transformation is post-transcriptional gene silencing (Johansen and Carrington, 2001). However, initial attempts to co-express with the p19

suppressor of gene silencing construct (Voinnet et al., 2003) failed to rescue expression (data not shown).

The most similar characterised homologue of *AtPQL1-3* is STM1, a putative GPCR from *S. pombe* (Chung et al., 2001). The vacuolar localisation of *AtPQL1* argues against a homologous function. Whereas GPCRs are localised to the cell surface, permitting binding of external ligands by an N-terminal facing the extracellular space, the vacuolar localisation of *AtPQL1* means it is likely that the N-terminal faces the vacuolar lumen. However, this does not necessarily rule out interaction with the sole *A. thaliana* Gα subunit, GPA1. *AtPQL1* and GPA1 could interact at points of contact between the tonoplast and plasma membrane. Precedents include THF1, a plastid localised protein which interacts with GPA1 at sites where the plastid membrane presses against the plasma membranes (Huang et al., 2006). In addition it cannot be ruled out that *AtPQL1* is present in both the tonoplast and plasma membranes.

If *AtPQL1-3*, like STM1, binds Gα subunits it would be expected that there would be overlap in tissue-specific localisation with GPA1. *GPA1* has been reported to be expressed throughout development in almost all organs examined with the exception of mature seeds and mature pollen (Weiss et al., 1993). However, GPA1 is found in both developing pollen and growing pollen tubes. Such an expression pattern does not exclude interaction with *AtPQL1* which is expressed almost exclusively in pollen, particularly developing pollen. However, it makes an interaction with *AtPQL2*, which is most strongly expressed in mature pollen unlikely.

4.3.1.2 Sub-group 2: *AtPQL4* and *AtPQL6*

Both the tissue-specific expression and sub-cellular localisation of *AtPQL4* and *AtPQL6* are consistent with them playing a similar role to their closest characterised homologue, MPDU1/LEC35. MPDU1/LEC35 is crucial for the utilisation of Man/Glc-P-Dol as sugar donors in glycosylation (Anand et al., 2001).

Firstly, it is known that the two best characterised forms of Man/Glc-P-Dol dependent glycosylation, N-glycosylation and GPI anchoring, occur in several different tissue types. Genes involved in the biosynthesis of N-glycans and/or GPI anchors

including *SETH1* and *SETH2* (Lalanne et al., 2004), *CYT1* (Lukowitz et al., 2001), and *GPT1* (Koizumi et al., 1999) were shown to be expressed in all organs tested. In addition, genes involved in the ER trimming and protein transfer of N-glycans including *GCSI* (Boisson et al., 2001) *RSW3* (Burn et al., 2002) and *DAD1* and *DAD2* (Danon et al., 2004) have been reported to be expressed widely.

In agreement with a potential role in glycosylation, both *AtPQL4* and *AtPQL6* were detected in all tissues tested. However, some small variations in expression were seen with *AtPQL4* showing slightly higher expression in open flowers and *AtPQL6* showing slightly higher expression in floral buds.

Secondly, it is known that MPDU1 acts at the ER membrane, where Man/Glc-P-Dol is synthesised and utilised (Anand et al., 2001). Although there are no reports of direct experimental localisation of MPDU1, or its CHO cell homologue LEC35, it is assumed these proteins localise to the ER (Ware and Lehrman, 1996; Anand et al., 2001). The fluorescence patterns produced by *AtPQL4*-GFP and *AtPQL6*-GFP fusion proteins are consistent with localisation to the ER. However, to categorically prove ER localisation, the proteins should be co-localised with an ER marker either by microscopy or biochemical methods.

4.3.1.3 Sub-group 3: *AtPQL5*

AtPQL5 has a notably stronger level of expression than the other *AtPQL* genes. *AtPQL5* transcript was found in all tissues tested but was particularly strongly expressed in open flowers, which is consistent with microarray data suggesting strong expression in mature pollen. Its overall higher level of expression may reflect the fact that it does not cluster with other *AtPQL* proteins in phylogenies of the PQL family (see chapter 3, figure 3-4). If it plays a distinct function and is the sole protein acting in such a role, it may be necessary for *AtPQL5* to be expressed at a higher level.

AtPQL5, like *AtPQL1*, appears to be localised to the tonoplast. This is consistent with a function in common with the closest characterised homologue of *AtPQL5*, CTNS. In humans, CTNS is localised to the lysosomes and exports cystine produced as a by-product of lysosomal protein hydrolysis (Cherqui et al., 2001; Kalatzis et al., 2001).

Given that the vacuole is the primary location of protein hydrolysis in plants (Callis, 1995) it would be expected that a functional homologue of CTNS would be located at the tonoplast. In agreement with the results shown here, *AtPQL5* was recently identified in a proteomics study to identify tonoplast proteins (Jaquinod et al., 2007).

Given the widespread occurrence of protein hydrolysis in plants (Callis, 1995), it would be expected that a CTNS-like protein would be expressed ubiquitously. Consistent with this, *AtPQL5* transcript was detected in all tissues tested. However, *AtPQL5* does appear to be more strongly expressed in open flowers. This may reflect the high rate of protein turnover known to be associated with programmed cell death during tapetal disintegration prior to anther dehiscence (Vierstra, 1993). However, microarray studies indicate that the highest expression is in mature pollen grains rather than sporophytic parts of the flower.

4.3.2 Common features amongst family members

4.3.2.1 Overlapping tissue-specific localisation

To an extent, there is overlap in the tissue-specific expression profiles of *AtPQL1-6*. All six transcripts were detected in all four tissues tested (negative controls gave no signal). However, all family members showed variation between organs. Figure 4-13 shows the relative expression level of each family member in the four tissue types tested. All but one showed peak expression in floral buds or open flowers. The exception, *AtPQL3* was most strongly expressed in roots.

Of the five genes most strongly expressed in flowers, three (*AtPQL1*, *AtPQL2* and *AtPQL5*) appear from microarray and promoter::GUS data to be preferentially expressed in pollen. Microarray data shows that another, *AtPQL6* is also expressed in pollen. It was not possible to check pollen-specific microarray data for the other two family members, *AtPQL3* and *AtPQL4*, as they are not represented on the ATH1 array but both showed expression in floral tissue.

Such overlap in expression raises the possibility of functional redundancy between family members, particularly those most closely related (i.e. *AtPQL1* and

AtPQL2, *AtPQL4* and *AtPQL6*). However, closer observation of microarray data shows that they are expressed at different times during pollen development.

Whereas *AtPQL1* and *AtPQL6* are most strongly expressed in developing pollen, *AtPQL2* and *AtPQL5* show higher expression in mature pollen. This is reflected in the qPCR results presented here, with *AtPQL1* and *AtPQL6* showing peak expression in floral buds (when the flowers contain more developing pollen) and *AtPQL2* and *AtPQL5* showing highest expression in open flowers (when most pollen has matured). It is possible that this reflects temporal segregation of function. For example, *AtPQL1* and *AtPQL2* may play a similar function but be specialised to act at different stages of pollen development.

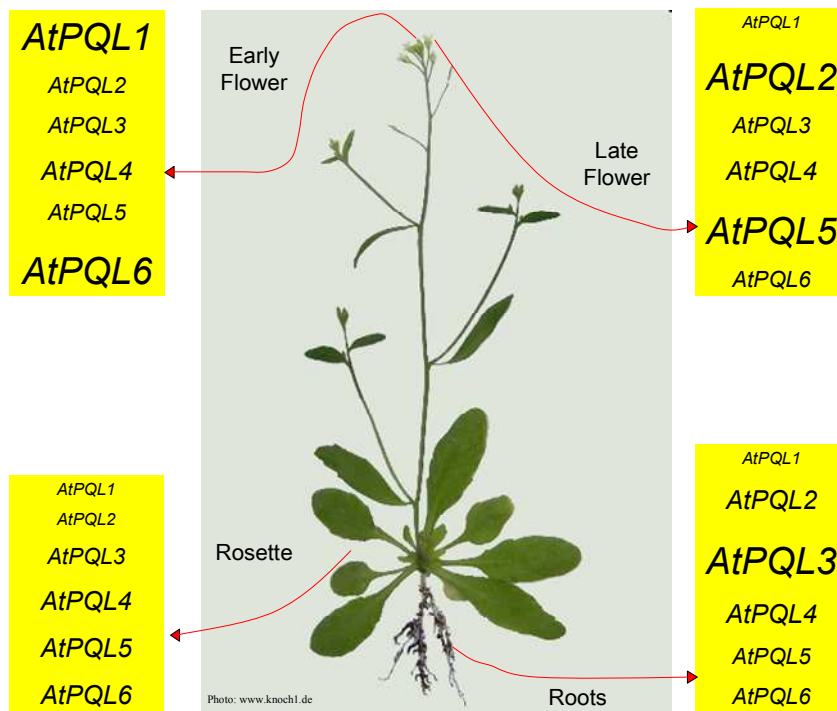


Figure 4-13. Relative tissue-specific expression level of *AtPQL1-6*. Variations in the relative expression level for each individual gene is reflected by changes in font size. The larger the font, the stronger the expression level for that gene in that tissue relative to the other tissues. Photo: www.knochl.de

4.3.2.2 Response to viral infection and wounding

Interestingly, a common feature of four *AtPQL* genes was induction in leaves by viral infection and wounding, especially in the *coil-16* background. Such regulation was

found for *AtPQL3*, *AtPQL4*, *AtPQL5* and *AtPQL6* (*AtPQL1* and *AtPQL2* were not checked due to their very low level of expression in vegetative tissues). The response of *AtPQL3* was particularly interesting. Whereas *AtPQL4*, *AtPQL5* and *AtPQL6* responded in wild-type plants but showed an amplified response in *coil-16*, *AtPQL3* responded only in *coil-16* and not at all in wild type. Although this data must be treated with caution due to the lack of biological replicates, the fact that similar results were seen for multiple genes in response to two independently applied treatments is encouraging.

The common regulation of a multi-gene family is not unusual. For example, all 14 members of the *HSP70* family are induced by heat-shock (Sung et al., 2001). It may reflect a common function of the *AtPQL* proteins in the response to viral infection and wounding (see chapter 6).

On the other hand it may be that each gene plays distinct functions but forms part of a common response. For example, *AtPQL4* and *AtPQL6* may have been induced because of an increased requirement for Man/Glc-P-Dol dependent glycosylation in response to virus infection and wounding. Interestingly, there is an increasing amount of evidence linking the secretory pathway (and glycosylation) with defence (see chapter 6).

The fact that the response was amplified (or in the case of *AtPQL3*, present exclusively) in the *coil-16* background is interesting. Two distinct possibilities can be envisaged. Firstly, the JA insensitivity of *coil-16* may have, in effect, led to greater stress severity, as these plants are unable to mount the normal JA response. Consistent with this hypothesis, no increase in *AtPQL* expression level was seen in the absence of stress. The second possibility is that regulation of the *AtPQL* genes is directly influenced by JA signalling. According to this hypothesis COI1 (or a downstream signalling component) negatively regulates *AtPQL* expression during stress. The absence of this negative regulator could explain the enhanced response in *coil*.

4.3.2.2 Diurnal changes in expression

Three family members showed changes in expression over the course of a day (*AtPQL4*, *AtPQL5* and *AtPQL6*). *AtPQL4* and *AtPQL6* both showed increasing expression levels throughout the day reaching a peak at the end of day around the time

the growth room lights were switched off. One of these genes, *AtPQL6* contains both an ‘evening element’ promoter motif and a CCA1 binding site in its promoter region. Both promote expression towards the end of day (see chapter 3) which is consistent with the results shown here.

AtPQL5 showed peak expression in the middle of the day in short-day conditions. In long days, expression increased in the middle of the day before decreasing and then peaking again around the times the lights were switched off. Although, more replicates would be required to determine whether this dip at 8 p.m. and increase at midnight is significant a similar pattern is seen in published microarray data.

For *AtPQL4* and *AtPQL6*, which are predicted to be involved in Man/Glc-P-Dol utilisation, their diurnal expression pattern may reflect changes in demand for glycosylation. To see if other genes known or predicted to be involved in Man/Glc-P-Dol dependent glycosylation (list of genes generated using KEGG database; Kanehisa et al. 2000) were regulated diurnally, their expression was checked using the *DIURNAL* microarray viewer (Mockler, 2007). Interestingly, *At3g57220*, one of two genes encoding homologues of GPT, the enzyme which catalyses the first step in N-glycan biosynthesis - the transfer of GlcNAc-1-P to Dol-P (see chapter 1, figure 1-2, step 1), shows a similar pattern of expression to *AtPQL6* in short days (see figure 4-14). In CHO cells, a genetic interaction has been established between *GPT* and *LEC35*, the CHO homologue of *AtPQL4* and *AtPQL6* (Gao et al., 2007a). It was found that over-expression of GPT directly interfered with the function of LEC35, albeit by an unknown mechanism (see chapter 1 section 1.3.7.4). If the same is true in plants, the similar diurnal regulation of these genes may reflect the importance of this interaction and the genes may be regulated in such a way so that GPT does not adversely affect *AtPQL4* and *AtPQL6* function.

In addition there was striking similarity in the expression pattern during long days between *AtPQL6* and *AtIg78800*, a gene encoding a homologue of ALG1 and ALG2, the enzymes responsible for the second and third mannosylation steps during N-glycan biosynthesis (see figure 4-14). This may reflect co-regulation in accordance with changes in demand for N-glycosylation.

For *AtPQL5*, whose closest characterised relative is CTNS, a mammalian lysosomal cystine exporter, diurnal changes in expression may reflect regulation of sulfur

assimilation and metabolism. Cystine is the disulfide form of the key sulfur metabolite cys. The two forms are readily interchangeable and following export to the cytosol by CTNS, cystine is rapidly reduced to cys (Cooper, 1983). Although it is unknown whether cystine export to the cytosol significantly contributes to the cytoplasmic cys pool in plants, it is known that in animals CTNS is required to maintain cys levels for normal rates of glutathione synthesis (Chol et al., 2004).

In plants, sulfur assimilation and metabolism is regulated diurnally and responds to light (Saito, 2004). It is generally thought that a group of circadian regulated genes ensure a large cytoplasmic pool of cys at the end of the night in preparation for an increase in sulphur metabolism when ATP, electrons, and 3-phosphoglycerate become available at the onset of photosynthesis in the morning (Saito, 2004). Given that *AtPQL5* is most strongly expressed in the middle of the day, this suggests it does not contribute to the accumulation of cys at the end of the night. However, it may be used to replenish cytoplasmic cys after depletion in the morning.

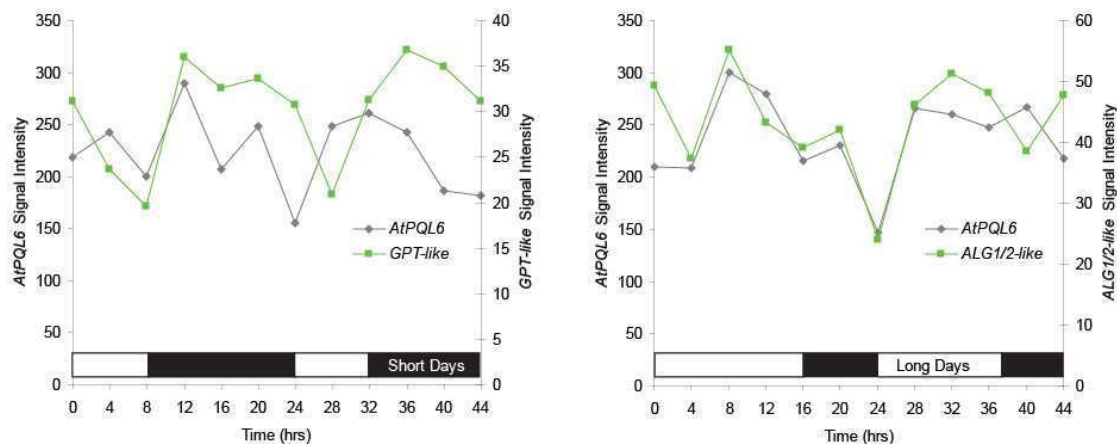


Figure 4-14. Microarray expression data for *AtPQL6*, and *ALG1/ALG2*- and *GPT*-like genes over the course of two days. Data was obtained from microarray experiments collated at the DIURNAL database (Mockler, 2007). The experiments presented used RNA extracted from *Ler0* seedlings grown on agar plates with Murashige and Skoog medium and 3% sucrose. Seedlings were grown in either long-day (16hr light/ 8hr dark) or short day (8hr light/16hr dark) conditions for 7 days before harvesting commenced. Plants were harvested at four hour intervals over the course of 48 hours. Data is reported as absolute signal intensity values and represents a single replicate for each time point.

Chapter 5: Functional Characterisation of the *AtPQL* Family

5.1 Introduction

5.1.1 Chapter organisation

In this chapter I will describe the functional characterisation of the *AtPQL* family. This chapter is divided into three parts; one for each of the different approaches used. These are heterologous expression in yeast (5.2), isolation and physiological characterisation of mutant lines (5.3), and transcriptional profiling of mutant lines (5.4). In this introduction I will explain the experimental strategy and rationale behind each approach.

5.1.2 Heterologous expression in yeast

Heterologous expression of plant genes in yeast is a well-established technique, and has been particularly successful in the characterisation of membrane proteins through functional complementation of yeast mutants. Recent examples of yeast mutants that have been rescued by plant genes encoding membrane proteins include the pH sensitive *kha1Δ* mutant by the *A. thaliana* cation/proton antiporter *AtCHX17* (Maresova and Sychrova, 2006); the slow growth phenotype of the *atm1Δ* mutants by the *A. thaliana* ATP-binding cassette transporter *AtATM3* (Chen et al., 2007) and the salt sensitive phenotype of the *enal-4Δ* strain by the rice potassium channel *OsKAT1* (Obata et al., 2007).

Even when no obvious candidate for functional complementation is available, expression in a heterologous system can yield valuable information through the study of dominant traits. For example, the expression of the *A. thaliana* calcium transporter, *CAX1* in tobacco resulted in an apical burning phenotype not seen in *A. thaliana* plants over-expressing the same gene (Hirschi, 1999). The relative ease and speed at which yeast can be screened for phenotypes makes it an ideal organism in which to express genes heterologously and look for dominant phenotypes.

Here, I examined the ability of *AtPQL* genes to complement the *ers1Δ* yeast strain. This strain carries a deletion of the yeast PQL family member *ERS1*, and has been reported to have a hygromycin sensitive phenotype that can be complemented by expression of the human PQL gene *CTNS* (Gao et al., 2005). I also checked for dominant traits associated with the heterologous expression of *AtPQL* family members in yeast. The outcome of this work is described in section 5.2.

5.1.3 Isolation and physiological characterisation of mutants

To assess the function of *AtPQL* proteins *in planta* a variety of different *A. thaliana* mutant lines were produced in this study. Knockout lines were isolated for all but one *AtPQL* family member and over-expresser mutants isolated for each.

5.1.3.1 Isolation of mutant lines

The knockout lines used in this study were from T-DNA insertion collections. Several different collections of T-DNA mutants are available for *A. thaliana* and together allow the selection of null mutations in almost any gene in the genome (Samson et al., 2002; Sessions et al., 2002; Rosso et al., 2003). Advantages include the high stability of the mutating DNA fragment, the inheritable nature of the T-DNA and the ability to mutate individual members of a gene family regardless of the level of homology within the family (Thornycroft et al., 2001).

In addition to knockouts, mutant lines over-expressing individual *AtPQL* genes were produced. Over-expression can be particularly useful when a gene is part of a gene family and functional redundancy prevents loss-of-function phenotypes. Accordingly, phenotypes are relatively common in gain-of-function mutants (Ichikawa et al., 2003).

In this study, the CaMV 35S promoter was chosen for over-expression of the *AtPQL* family (see figure 5-9). This promoter is widely used to drive constitutive expression and is active in almost all *A. thaliana* tissues with the notable exception of pollen (Jefferson et al., 1987; Wilkinson et al., 1997).

5.1.3.2 Physiological characterisation of mutant lines

A number of different experiments were conducted to look for physiological phenotypes. The rationale behind each experiment is described below.

Analysis of growth in normal conditions: Phenotypes based on organ size and timing of development are common in *A. thaliana* (Boyce et al., 2001). To check for such phenotypes, *AtPQL* mutants were grown under standard conditions on soil and agar plates.

Analysis of sensitivity to heat-shock: As well as looking under normal conditions, a number of experiments were carried out to test for conditional phenotypes. A common stress-sensitive phenotype is altered sensitivity to heat. A primary objective of the heat shock response is to maintain protein conformation and membrane stability (Wang et al., 2004). Given the proposed function of *AtPQL4* and *AtPQL6* in glycosylation and the importance of glycosylation in protein folding (see chapter 1 section 1.3.2.4), it was hypothesized that *AtPQL* mutants might display an altered response to heat.

Analysis of sensitivity to salt and dehydration: Another frequently reported mutant phenotype is altered salt sensitivity. Salinity causes both ionic and osmotic stress, resulting in impaired metabolism, disruption of membranes, reduced nutrient uptake, and an overall negative effect on plant fitness (Hasegawa et al., 2000).

One common aspect of salt and heat stress is that both disturb the stability of membranes and proteins (Wang et al., 2004). For the reasons mentioned above, *AtPQL4* and *AtPQL6* mutants may show altered salt tolerance as a result of glycosylation defects. In addition, plant survival under both salt stress and dehydration is dependent on the maintenance of ion homeostasis through the regulation of transport (Hasegawa et al., 2000). Given that the *AtPQL* family are membrane proteins one possible role is transport. If *AtPQL* proteins are involved in ion homeostasis, mutants may show a salt/drought dependent phenotype.

Analysis of sensitivity to cadmium, SNP and sulfur starvation: As one member of the PQL family from humans, CTNS, is a lysosomal cystine transporter (Kalatzis et al. 2001), a number of experiments were designed to test for phenotypes which could be explained by defects in cystine transport.

Firstly, some mutant lines were checked for altered cadmium sensitivity. The rationale being that cystine transport may be important in cadmium detoxification given the role of the reduced form of cystine - cysteine (cys) - as a precursor to glutathione (GSH). Studies have shown that the availability of cys directly influences GSH production (Noctor et al., 1996). GSH itself, is a precursor in the synthesis of phytochelatins which buffer cytosolic cadmium concentration by chelating heavy metals before being sequestered into the vacuole (Cobbett, 2000). A reduction in cys availability, as a result of a failure to export cystine from intracellular stores, might limit phytochelatin production and lead to cadmium hypersensitivity.

Secondly, the effect of sodium nitroprusside (SNP), a nitric oxide (NO) donor, was tested on some mutant lines. NO is produced by plants as a signalling component in a wide range of processes (Neill et al., 2002). At low concentrations, exogenously applied SNP has been reported to promote vegetative growth in *A. thaliana*, but at higher concentrations (above 100 μ M) has a toxic effect (He et al., 2004). NO has also been reported to induce GSH synthesis in roots and can react with GSH to form S-nitrosogluthathione (GSNO); a suggested long-distance transportable source of NO (Neill et al., 2002; Innocenti et al., 2007). A reduction in cys availability may limit GSH production and influence NO signalling.

Any cadmium/SNP dependent phenotype would presumably be more severe if cys was limited. Levels of cys, as the major sulfur containing amino acid, are reduced during sulfur starvation (Nikiforova et al., 2003). In order to check for an amplified effect due to limited cys, experiments were carried out using control and low sulfur media.

Analysis of AtPQL mutant pollen: As some *AtPQL* genes are preferentially expressed in pollen (*AtPQL1*, *AtPQL2* and *AtPQL5*; see chapter 4), it was hypothesized that mutant lines may show defects in fertility. The fact that homozygous lines could be found, proves that none are completely sterile. However, changes in pollen characteristics have been reported even in viable mutants. For example, *gum* and *mud* mutants have been reported to show defects in the placement of pollen nuclei, but still retain viability albeit with reduced transmission (Lalanne and Twell, 2002). Here, possible pollen phenotypes were examined by testing pollen viability and germination *in vitro*.

5.1.4 Transcriptional profiling of mutant lines

5.1.4.1 Microarray experiments

The transcriptional profiles of four different mutant lines, *pql4-1*, *pql4-2*, *pql6-1* and *35S::AtPQL6-1*, were examined using the Affymetrix ATH1 full-genome microarray. The ATH1 chip, is a single channel oligonucleotide array, with 25-mer probes representing approximately 24,000 *A. thaliana* genes.

Transcriptional profiling has often proved useful for mutant characterisation, on some occasions helping to confirm an already suspected function. For example, transcriptional profiling of mutants with altered expression of a heat and light responsive transcription factor, identified changes in the expression of genes known to be involved in these processes (Nishizawa et al., 2006). Transcriptional profiling can also shed light on additional (and sometimes surprising) roles for genes of known function. For example, transcriptional profiling of a knockout mutant lacking NHX1, a sodium proton antiporter, showed that like its yeast homologue, NHX1 function was not restricted to salt homeostasis but was also involved in endosomal trafficking (Sottosanto et al., 2004).

In this study, normalised signals from the arrays were analysed using rank products (RP; Breitling et al., 2004b) and iterative group analysis (iGA; Breitling et al. 2004a). RP analysis ranks genes according to the rank of their fold changes in each individual replicate. Decisions about which genes are considered ‘differentially expressed’ can be based on calculated false detection rates (FDR), which represent the likelihood of a certain rank being obtained by chance. iGA tests for differential expression of groups of functionally related genes and can be useful for detecting changes that are not obvious when looking at individual genes. It is based on the RP lists and identifies groups of genes which show a distribution biased towards the top of the list. One of its main advantages is that it does not rely on pre-defined lists of differentially regulated genes, thereby being able to detect tendencies within expression profiles that are easily overlooked if only a small part of the dataset is used for functional interpretation. Both methods have been used successfully in many microarray studies (Nemhauser et al., 2006; Keurentjes et al., 2007). Details of both are given in materials and methods (see section 2.5.7.2).

5.1.4.2 Links between glycosylation and gene expression

The mutants used for microarray experiments have altered expression of either *AtPQL4* or *AtPQL6*, two genes that are homologous to *LEC35/MPDU1* and predicted to play a role in glycosylation. There are several examples of links between N-glycosylation and gene expression. Three previous studies have profiled *A. thaliana* gene expression in response to tunicamycin (an inhibitor of the first step of N-linked glycosylation) to investigate ER stress and the unfolded protein response (UPR). Two studies used partial genome arrays (c. 8200 probes) and found up-regulation of several UPR-associated genes encoding ER-localised protein folding chaperones and trafficking related genes (Martinez and Chrispeels, 2003; Noh et al., 2003). A more recent study, using fluid microarrays (a competitive hybridisation approach where DNA fragments are cloned onto microbeads and hybridised with differentially labelled probes from treated and non-treated plants) also found up-regulation of several ER chaperones as well as genes involved in glycosylation and vesicle trafficking (Kamauchi et al., 2005). There was considerable overlap in the genes identified. Common up-regulated genes encoding ER-folding chaperones (such as BiP, calnexin and calreticulin), UTr1 (see chapter 1; section 1.3.2.4), and proteins involved in trafficking and degradation were found in all three studies.

5.2 Results: Heterologous expression in yeast

5.2.1 Phenotype of *ers1Δ* deletion mutants

The *ers1Δ* deletion mutant has been reported to have a hygromycin B sensitive phenotype (Gao et al., 2005). Hygromycin is an antibiotic that inhibits protein synthesis (Gonzalez et al., 1978). Hygromycin tolerance of yeast is linked to a number of physiological functions (see section 5.5.1). To assess hygromycin sensitivity, the yeast was grown to exponential phase in liquid medium and subsequently plated out on agar-solidified media containing different concentrations of hygromycin B. Droplets of decreasing culture density were applied in rows, while different strains were placed next to each other in columns (yeast drop assay). The strains used are listed in table 5-1. *ers1Δ*

cells were transformed with either an integrative plasmid containing the wild-type *ERS1* or *CTNS* gene, or an episomal plasmid containing an *AtPQL* gene, *ERS1* or no insert. As can be seen in figure 5-1, the hygromycin sensitive phenotype of the *ers1Δ* strain could not be reproduced. No difference in growth between the *ers1Δ* strain and the *ers1Δ* strain transformed with wild-type *ERS1* (*ers1Δ::ERS1*), could be seen at any of the hygromycin concentrations tested. The same was true when the experiment was carried out with synthetic complete or YPAD medium instead of YPD (data not shown).

Although no difference in hygromycin sensitivity was seen between *ers1Δ* and *ers1Δ::ERS1*, yeast transformed with *CTNS* (*ers1Δ::CTNS*) showed slightly better growth than *ers1Δ* on hygromycin (at 30μg/mL and above). As no reproducible hygromycin sensitive phenotype could be found, a screen was conducted to search for alternative phenotypes of *ers1Δ*. If a second phenotype could be found this would allow for testing of functional complementation by *AtPQL* genes. As shown in figure 5-2, a variety of different conditions were tested for effects on growth. This included supplementing the growth medium with ethanol, sorbitol, NaCl, LiCl, CdCl₂, CaCl₂, ZnCl₂, MgCl₂ or CuSO₄; incubating at different temperatures (17°C and 37°C) or pH (pH 3.5); and diluting synthetic medium to 1/2, 1/10 or 1/100 of the normal strength. In these experiments the strains tested were *ers1Δ*, *ers1Δ::ERS1*, *ers1Δ::CTNS* and *ers1Δ* transformed with the empty episomal vector pDR195 (*ers1Δ* + pDR195 EV). Unfortunately no alternative phenotypes could be found for *ers1Δ*.

Table 5-1. Yeast strains used in this study.

Name	Genotype	
<i>ers1Δ</i>	<i>ers1Δ::His3</i>	

	Plasmid Used For Transformation	Plasmid Type
<i>ers1Δ::ERS1</i>	pRS305 <i>P_{ERS1}::3xMyc-ERS1</i>	Yip
<i>ers1Δ::CTNS</i>	pRS305 <i>P_{ERS1}::CTNS</i>	Yip
<i>ers1Δ</i> + pDR195 EV	pDR195 <i>P_{PMA1}::empty vector (no insert)</i>	Yep
<i>ers1Δ</i> + pDR195 <i>AtPQL1</i>	pDR195 <i>P_{PMA1}::AtPQL1</i>	Yep
<i>ers1Δ</i> + pDR195 <i>AtPQL2</i>	pDR195 <i>P_{PMA1}::AtPQL2</i>	Yep
<i>ers1Δ</i> + pDR195 <i>AtPQL4</i>	pDR195 <i>P_{PMA1}::AtPQL4</i>	Yep
<i>ers1Δ</i> + pDR195 <i>AtPQL5</i>	pDR195 <i>P_{PMA1}::AtPQL5</i>	Yep
<i>ers1Δ</i> + pDR195 <i>AtPQL6</i>	pDR195 <i>P_{PMA1}::AtPQL6</i>	Yep

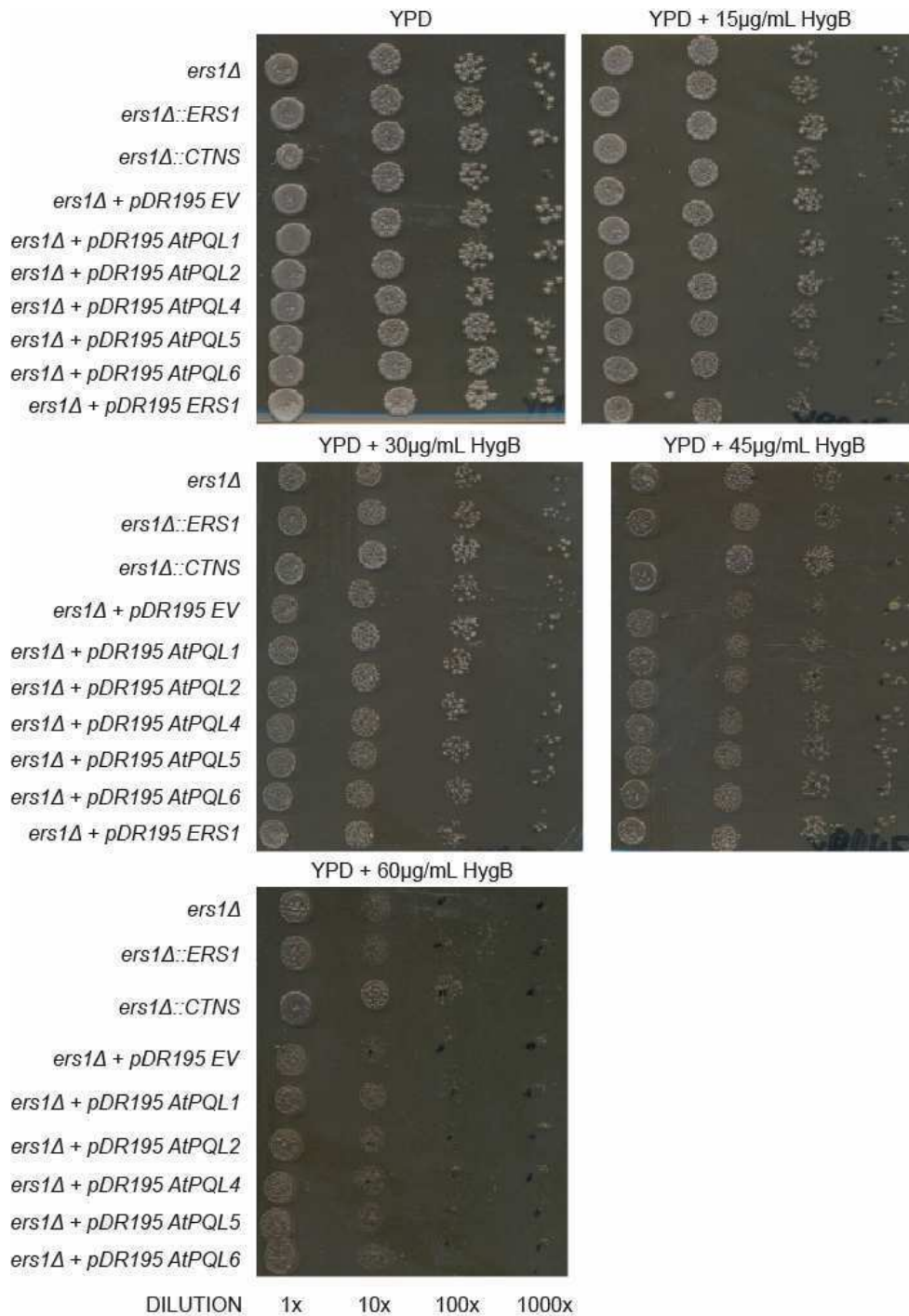


Figure 5-1. Growth of yeast strains on hygromycin. *ers1Δ* yeast was transformed with genes encoding PQL proteins and applied to YPD medium (2% yeast extract, 4% peptone, 4% dextrose, pH7) containing various concentrations of hygromycin. Yeast strains growing in exponential phase were applied to the medium at an OD₆₀₀ of 0.2 and in three ten-fold serial dilutions (left to right). Each plate was then incubated at 30°C for two days.

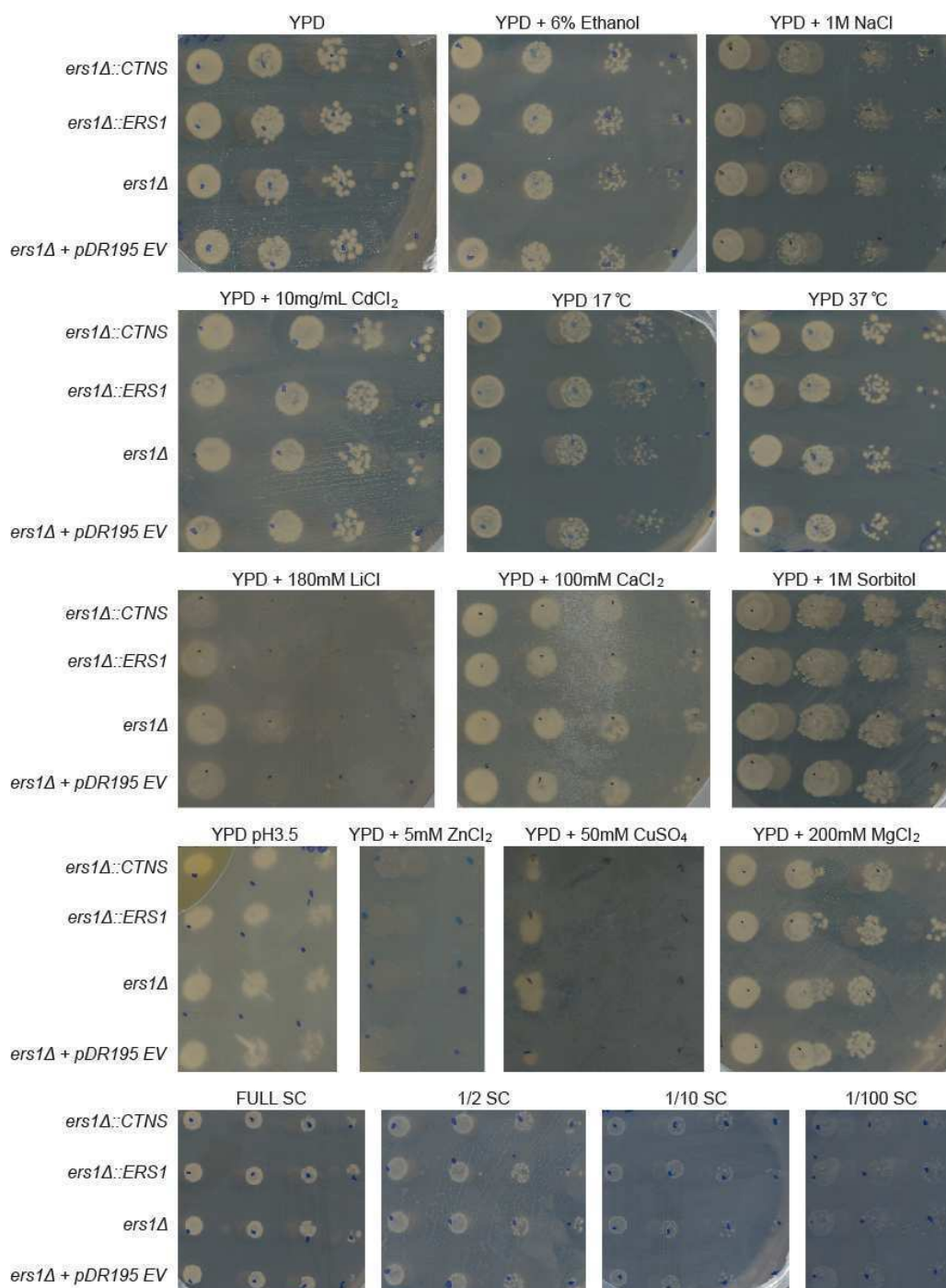


Figure 5-2. Growth of yeast strains under a variety of conditions. The indicated yeast strains were grown under the stated conditions. Growth and background conditions were the same as for figure 5-1. Except where indicated otherwise the medium was adjusted to pH7 and the plates incubated at 30°C for two days (four days for the 17°C treatments). Full SC medium consisted of 0.34% yeast nitrogen base, 1% ammonium sulphate, 2% glucose and a complete supplement mix of amino acids as described in table 2-6.

5.2.2 Expression of *AtPQL* proteins in yeast

Due to the lack of phenotype in *ers1Δ* it was not possible to check for functional complementation with *AtPQL* genes. Nevertheless, differential growth in yeast transformed with *AtPQL* genes could still give clues as to function. Open reading frames of the *AtPQL* genes were cloned into the yeast episomal vector pDR195 under the control of the *PMAl* promoter (see table 5-1). The plasmids were then transformed into *ers1Δ* and the resulting strains used in growth assays. For one family member, *AtPQL3*, cloning into this expression vector failed despite several attempts.

None of the expressed *AtPQL* genes had an effect on the hygromycin sensitivity of *ers1Δ* (see figure 5-1). Since there were some initial indications from transgenic plant lines of a glucose related phenotype (later found to be irreproducible), growth of the strains was also checked on synthetic medium containing various concentrations of glucose. As can be seen in figure 5-3, all strains showed very little growth on medium without glucose (although some growth still occurred due to residual glucose carried over from the liquid medium). Growth improved as glucose concentration was raised up to 5%. At concentrations higher than 5%, an inhibitory affect on growth was observed. No differences were observed between any of the strains tested.

Finally, given that one human PQL protein CTNS is a cystine transporter some of the strains were also checked for sensitivity to cystine. As can be seen in figure 5-4, the addition of cystine to the media had little effect on growth, and no differences were observed between strains.

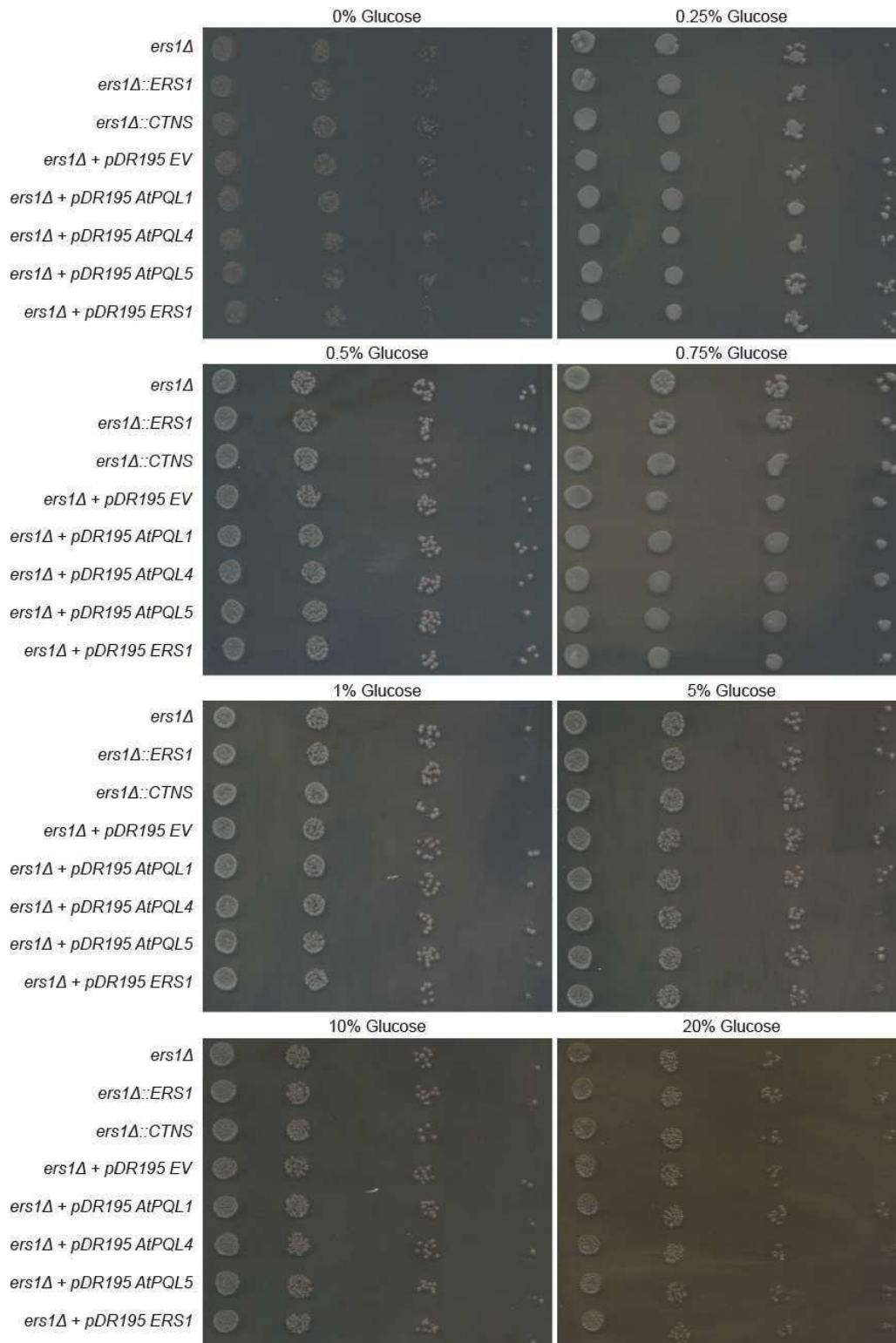


Figure 5-3. Growth of yeast strains on low and high glucose media. *ers1Δ* yeast was transformed with genes encoding PQL proteins and applied to synthetic complete medium with various concentrations of glucose. Growth and background conditions were the same as for figure 5-1. SC medium consisted of 0.34% yeast nitrogen base, 1% ammonium sulphate and a complete supplement mix of amino acids as described in table 2-6.

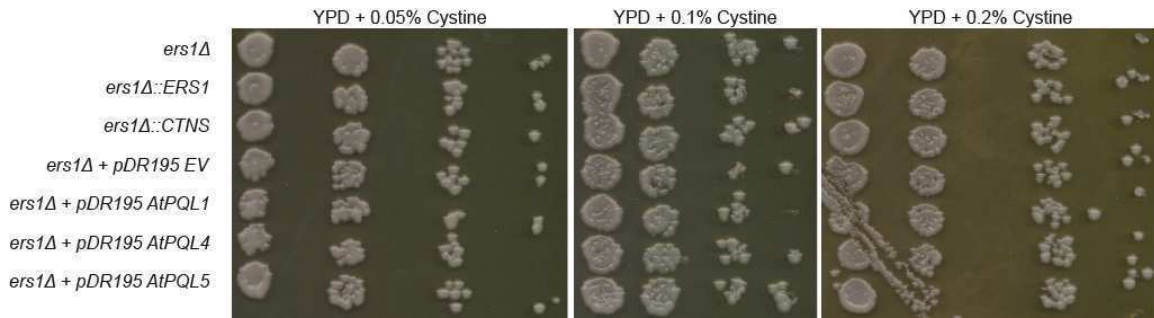


Figure 5-4. Growth of yeast strains on cystine-containing media. *ers1Δ* yeast was transformed with genes encoding PQL proteins and applied to YPD medium containing the indicated concentration of cystine. Growth and background conditions were the same as for figure 5-1.

5.3 Results: Isolation and physiological characterisation of mutants

5.3.1 Isolation of T-DNA insertion mutants

All T-DNA insertion mutants used in this study were obtained from the SALK collection. Figure 5-5 shows the insertion site for each line. To obtain homozygous lines, segregating T₃ or T₄ plants were genotyped by PCR according to the scheme outlined in figure 5-6a. The progeny of individual homozygous plants were used for subsequent characterisation. Results of diagnostic PCR experiments, proving each line is homozygous are shown in figure 5-6b-g. In each case PCR with primers spanning the insertion site (LP-RP) gave no amplification proving the wild-type allele for each gene was absent. PCR with a second set of primers, one from the left border of the T-DNA (Lba1) and one from the genomic sequence flanking the insertion site (RP), gave an amplification product proving that the T-DNA insertion was present.

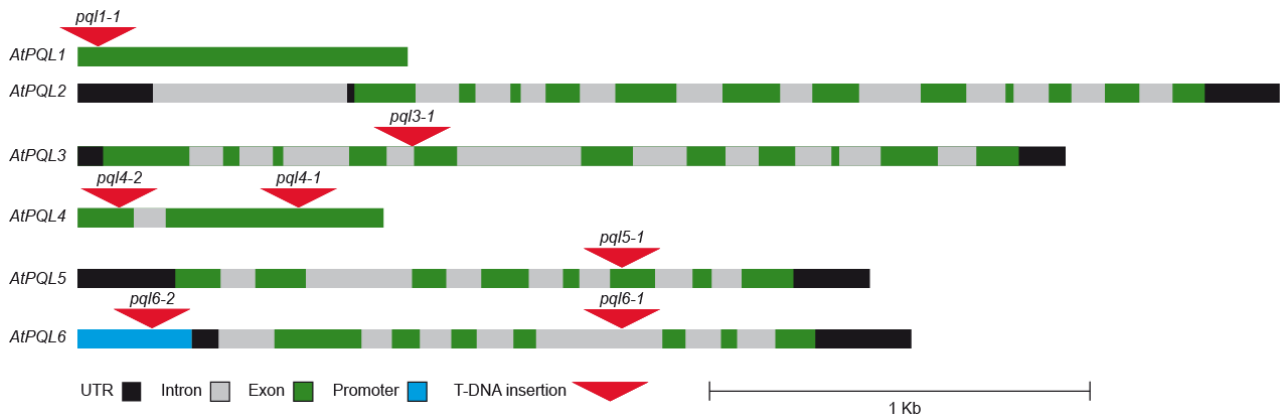
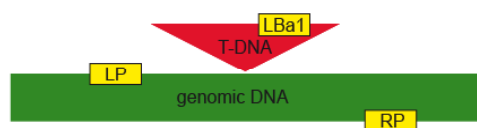
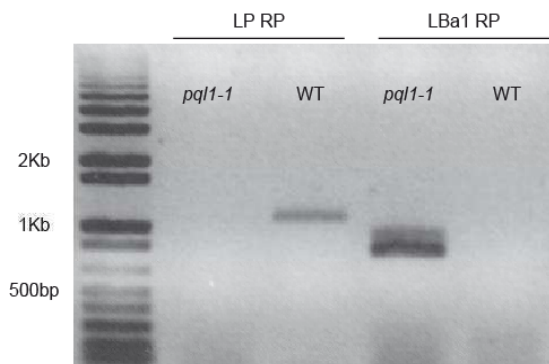


Figure 5-5. T-DNA insertion sites of SALK mutant lines used in this study.

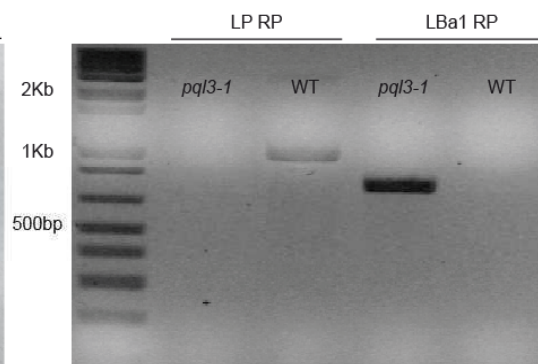
A. PCR Primer Location



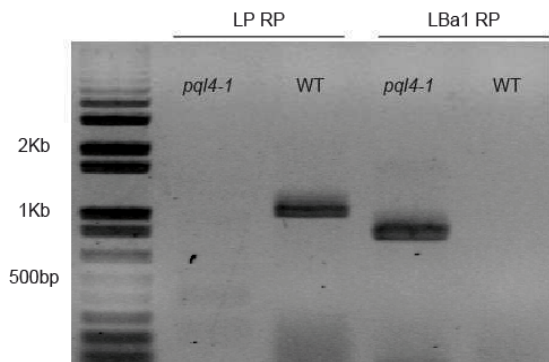
B. *pql1-1* SALK_108796



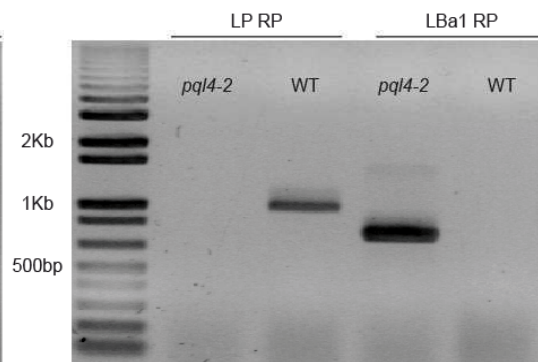
C. *pql3-1* SALK_044346



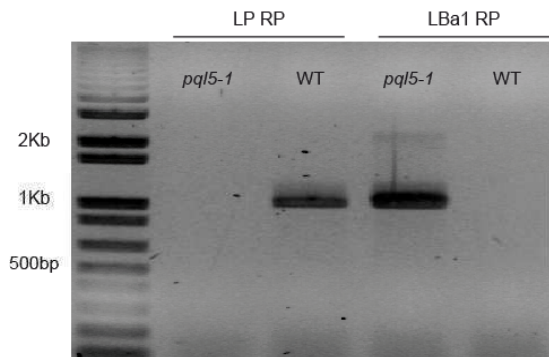
D. *pql4-1* SALK_018034



E. *pql4-2* SALK_000090



F. *pql5-1* SALK_118698



G. *pql6-1* SALK_151293

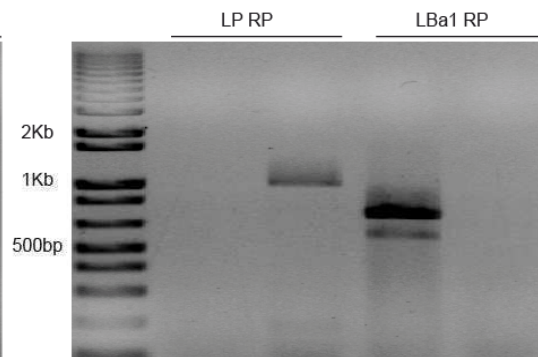


Figure 5-6. Confirmation of T-DNA knockout lines at the gDNA level. PCR was carried out for each line with two specific primer pairs to check for the presence of wild-type and T-DNA knockout alleles. **A.** Position of primers relative to T-DNA. LP and RP primers are specific for genomic DNA flanking the insertion site and the LBa1 primer is specific for the T-DNA. **B-G.** PCR products amplified from mutant and wild-type gDNA. In each case, a T-DNA allele-specific band was only amplified from mutant gDNA whereas a wild-type allele-specific band was only amplified from wild-type gDNA.

5.3.2 Genetic characteristics of T-DNA insertion mutants

The genetic characteristics of each insertion mutant are described below.

AtPQL1: One homozygous T-DNA insertion mutant was obtained for *AtPQL1*. *pql1-1* has an insertion 50 bp downstream of the start codon.

AtPQL2: *AtPQL2* is the only family member for which no insertion mutant was isolated. No T-DNA mutants are available with an insertion site within the coding region of *AtPQL2*.

AtPQL3: One line was found for *AtPQL3*. The insertion site in *pql3-1* is close to the centre of the gene at the 3' end of the fourth intron. As the T-DNA is present in an intron it is possible that it could be removed by splicing. However, no transcript was detectable by reverse-transcription PCR with primers spanning the insertion site (see figure 5-7a).

AtPQL4: Two homozygous knockout lines were isolated for *AtPQL4*. The first, *pql4-1* has an insertion in the second exon, 224 bp 5' of the stop codon. The second, *pql4-2* has an insertion in the first exon, 109 bp 3' of the start codon.

AtPQL5: One homozygous line was found for *AtPQL5*. *pql5-1* has an insertion towards the 3' end of the gene in the sixth of eight exons.

AtPQL6: For *AtPQL6*, one knockout line, *pql6-1*, was found in the fifth intron. To check if transcript was still produced due to splicing of the T-DNA, real-time reverse-transcription PCR was performed with primers 5' of the insertion site (see figure 5-7b). Although transcript was still produced, it was less than half as abundant. A second insertion line, *pql6-2* was found with an insertion site 100 bp upstream of the start codon. However, semi-quantitative reverse-transcription PCR showed no reduction in transcript compared to wild type (figure 5-7c).

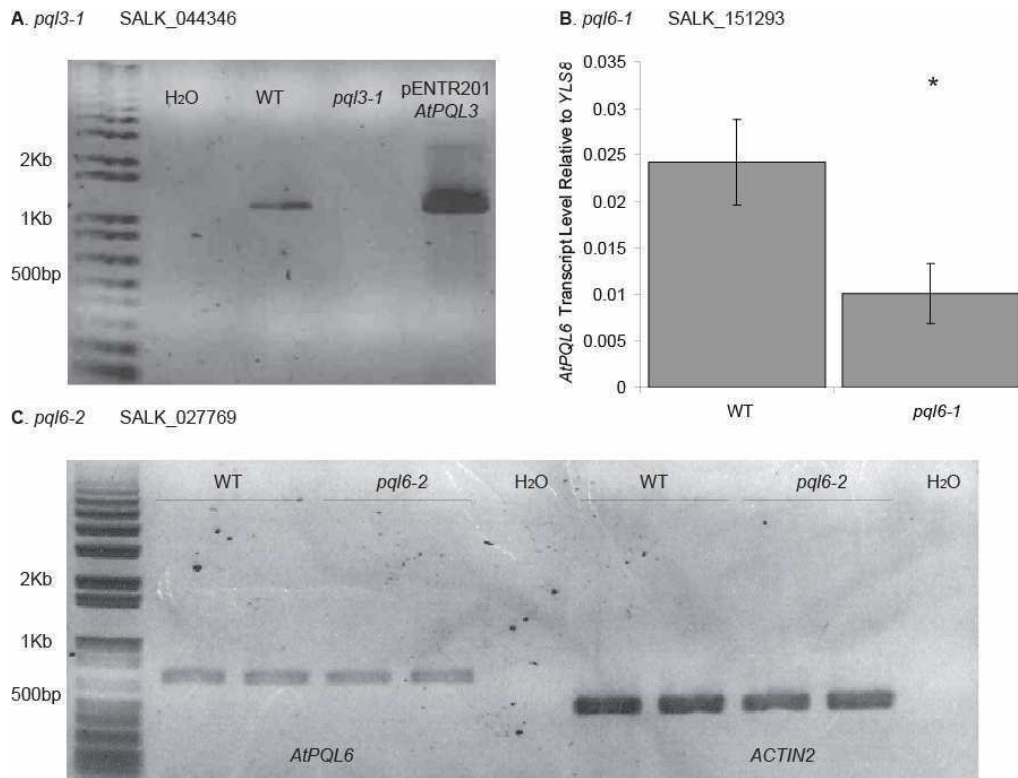


Figure 5-7. Confirmation of T-DNA knockout lines at the transcript level. **A.** confirmation of the absence of *AtPQL3* transcript in *pql3-1* by reverse-transcription PCR. cDNA from individual wild-type and *pql3-1* plants was used as starting material. Plasmid DNA is shown as a positive control. **B-C.** *AtPQL6* transcript level in *pql6-1* and *pql6-2*. **B.** *AtPQL6* transcript level was assessed in a pooled sample of 15 seedlings grown on agar plates for 12 days by real-time reverse-transcription PCR. Average transcript level, relative to *YLS8*, was calculated from three biological replicates. Error bars indicate standard error. A t-test was used to check for significant differences between wild type and *pql6-1*. * = $t < 0.1$. **C.** *AtPQL6* transcript level was assessed by semi-quantitative reverse-transcription PCR using cDNA from two individual wild-type and *pql6-2* plants as starting material. Amplification of a reference gene, *ACTIN2*, is also shown.



Figure 5-8. Expression cassettes of plasmids used in this chapter. The diagram shows the T-DNA cassette of the pGWB2 vector. In each case the Gateway attB recombination sites and plant selectable markers (HPTII and NPTII confer resistance to hygromycin and kanamycin respectively) are indicated. The position of the *AtPQL1-6* open reading frames relative to the CaMV 35S promoter and *NOS* terminator is shown.

5.3.3 Over-expresser mutant isolation

At least one over-expresser line was isolated for each *AtPQL* family member. Over-expression was driven by the cauliflower mosaic virus (CaMV) 35S promoter in the pGWB2 vector (Nakagawa et al., 2007). The features of these constructs are shown in figure 5-8.

Wild-type plants were transformed by floral-dip (Clough and Bent, 1998) and T₁ transformants selected by hygromycin resistance. To obtain homozygous lines with a single insert, the segregation of the hygromycin resistance marker was examined in the T₂ and T₃ generations (see table 5-2). In each case T₂ plants segregating 3:1 (indicating a single insert) and T₃ plants showing 100% resistance (indicating homozygosity) were used for further characterisation.

In each line checked, the transgene was successfully over-expressed. As shown in figure 5-9a, semi-quantitative reverse-transcription PCR confirmed that the expression of *AtPQL1* was increased in two independent lines relative to wild type whereas the expression of a reference gene remained stable. Similarly as shown in figures 5-9b and 5-10c quantitative PCR confirmed over-expression of *AtPQL4* and *AtPQL6* in two independent lines.

Table 5-2. Segregation analysis of *AtPQL* over-expressor lines. All transgenic lines produced in this study contain a hygromycin-resistance marker. The table shows the proportion of T₂ and T₃ seeds, harvested from individual plants that grew on agar plates containing 40 µM hygromycin to the four-leaf stage. The χ^2 statistic indicates that none of the lines showed a segregation ratio in the T₂ generation significantly different from the 3:1 ratio expected for a single insert. Close to 100 % growth in the T₃ generation indicates a homozygous line.

	T ₂				T ₃		
	Growth	No Growth	%	χ^2	Growth	No Growth	%
<i>35S::AtPQL1-1</i>	44	13	77.19	0.70	64	0	100
<i>35S::AtPQL1-2</i>	37	10	78.72	0.56	105	2	98.13
<i>35S::AtPQL2-1</i>	65	29	69.15	0.19	80	2	97.56
<i>35S::AtPQL2-2</i>	61	24	71.76	0.49	70	0	100
<i>35S::AtPQL3-1</i>	49	13	79.03	0.46	44	0	100
<i>35S::AtPQL3-2</i>	60	19	75.95	0.85	97	0	100
<i>35S::AtPQL4-1</i>	31	9	78	0.72	49	0	100
<i>35S::AtPQL4-GFP-1</i>	50	20	71	0.49	71	1	99
<i>35S::AtPQL5-1</i>	51	18	73.91	0.83	62	2	96.88
<i>35S::AtPQL6-1</i>	79	20	79.80	0.27	48	0	100
<i>35S::AtPQL6-2</i>	58	22	72.50	0.61	55	2	96.5

5.3.4 Phenotypic analysis of mutant lines

5.3.4.1 Analysis of growth in normal conditions

To check for phenotypes under non-stressed conditions, plants were grown on soil and agar plates. Figure 5-10 shows representative images of *AtPQL6* mutants. At all stages of development, from seed through to flowering plant, no reproducible differences were found between mutant and wild type. Similar observations were made for all the other mutant lines used in this study (images not shown to avoid repetition).

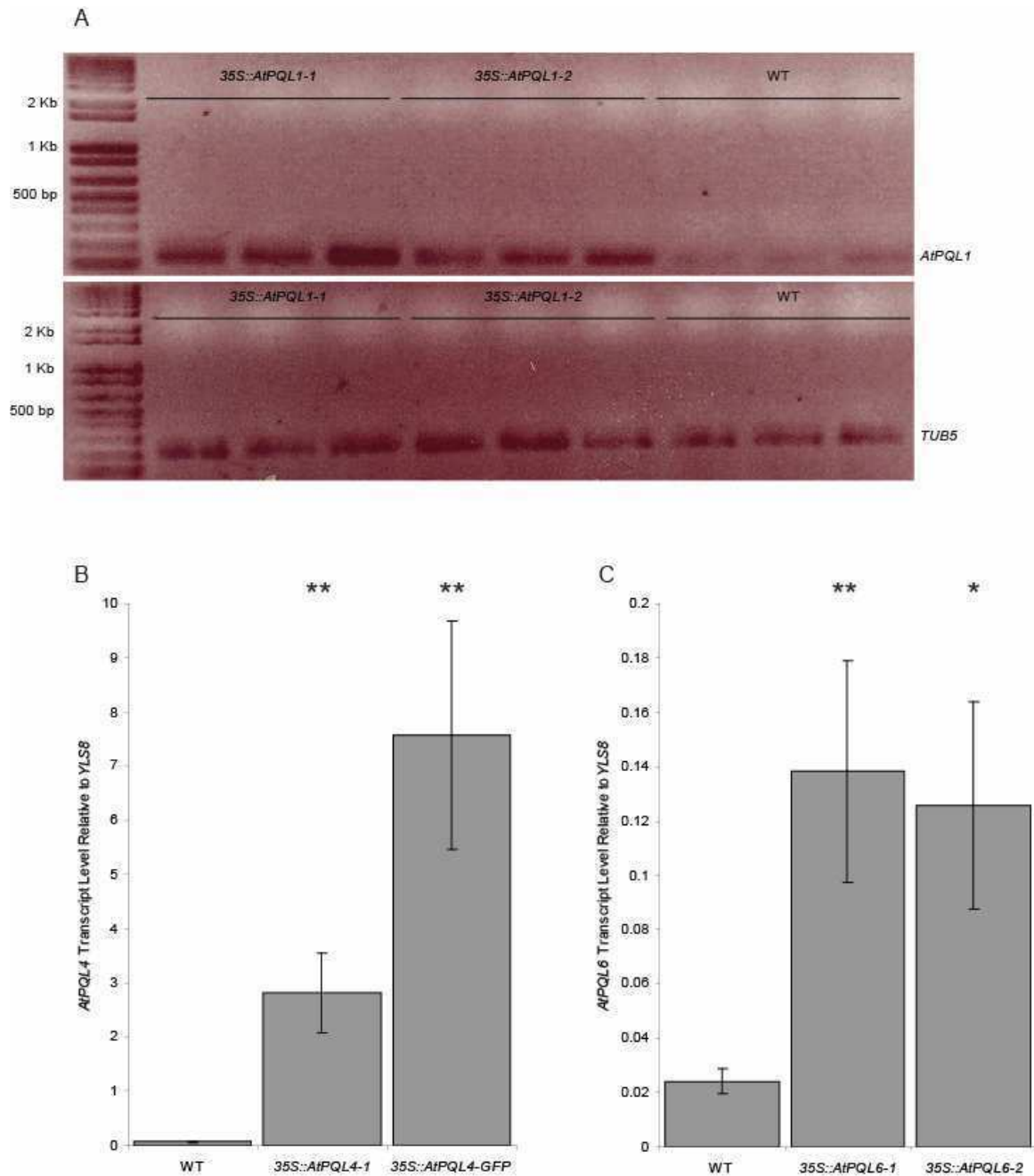


Figure 5-9. Confirmation of *AtPQL* over-expression. **A.** Semi-quantitative reverse-transcription PCR confirmation of *AtPQL1* over-expression in *35S::AtPQL1-1* and *35S::AtPQL1-2*. *AtPQL1* transcript was amplified from cDNA of three individual plants from each over-expressing line and wild type. Amplification of a reference gene, *TUB5*, was similar in all the plants tested. **B-C.** Confirmation of *AtPQL4* (**A**) and *AtPQL6* (**B**) over-expression by real-time reverse-transcription PCR. Transcript level was assessed in a pooled sample of 15 seedlings grown on agar plates for 12 days. Average transcript level, relative to *YLS8*, was calculated from three biological replicates. Error bars indicate standard error. A t-test was used to check for significant differences between wild type and each over-expressing line. * = $t < 0.1$ and ** = $t < 0.05$.

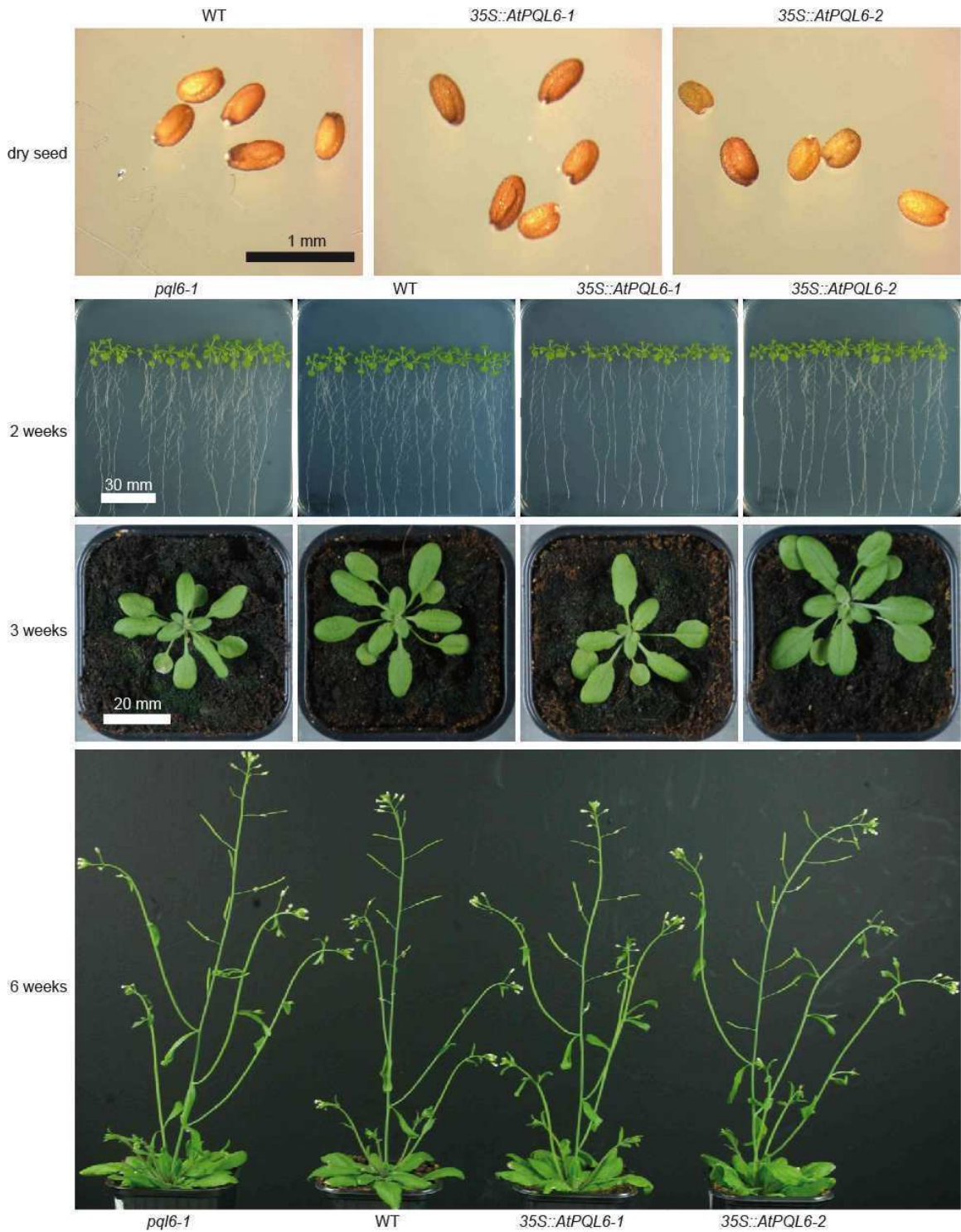


Figure 5-10. Images of wild type and *AtPQL6* mutant lines. Plants on agar plates were grown on control medium based on Arteca and Arteca (2000) with 3% sucrose. Photographs were taken 2 weeks after germination. Soil plants were grown in long-day conditions (16 hr light, 22 °C, 60% relative humidity / 8 hr dark, 18 °C, 70% relative humidity) and photographed after 3 or 6 weeks as indicated.

5.3.4.2 Sensitivity to heat shock

To check for differences in tolerance to high temperature, mutant lines were grown alongside wild type on soil in a temperature controlled growth room (21 °C) for three weeks before being heat shocked at 40 °C overnight. Images of *AtPQL4* and *AtPQL6* mutants before and after heat shock are shown in figure 5-11. (Again, images of other lines are not shown to avoid repetition)

The heat-shock treatment had a severe detrimental effect on all the lines tested. Immediately after the heat treatment leaves appeared wrinkled. After being returned to normal temperature for six days the majority of leaves appeared chlorotic and most plants did not recover. No differences in the severity of symptoms or likelihood of survival were found between wild type and any of the mutant lines tested.

5.3.4.3 Sensitivity to salt: soil-grown plants

To test salt sensitivity, each line was grown alongside wild type and watered with increasing concentrations of sodium chloride. Before the treatment, plants were watered regularly with normal watering solution (1/8 strength MS). After three weeks the watering solution was supplemented with 50 mM NaCl. The NaCl concentration was then gradually increased as described in the right-hand panel of figure 5-12.

The treatment had a progressive detrimental effect on each mutant line and wild type. For the first two weeks, the plants continued to grow and remained green. After two weeks, when the NaCl concentration reached 300 mM, symptoms of salt stress became obvious. No further growth was observed and chlorotic lesions appeared. By the end of the treatment (day 31) almost all plants were completely chlorotic.

No differences were seen in the severity or timing of symptoms in any of the lines tested. Although occasionally, green leaves were still visible at day 31, they did not seem to be more or less prevalent in any of the mutant lines compared to wild type. Similarly, some plants began to flower during the treatment but again there was no indication of differences in the proportion of flowering plants.

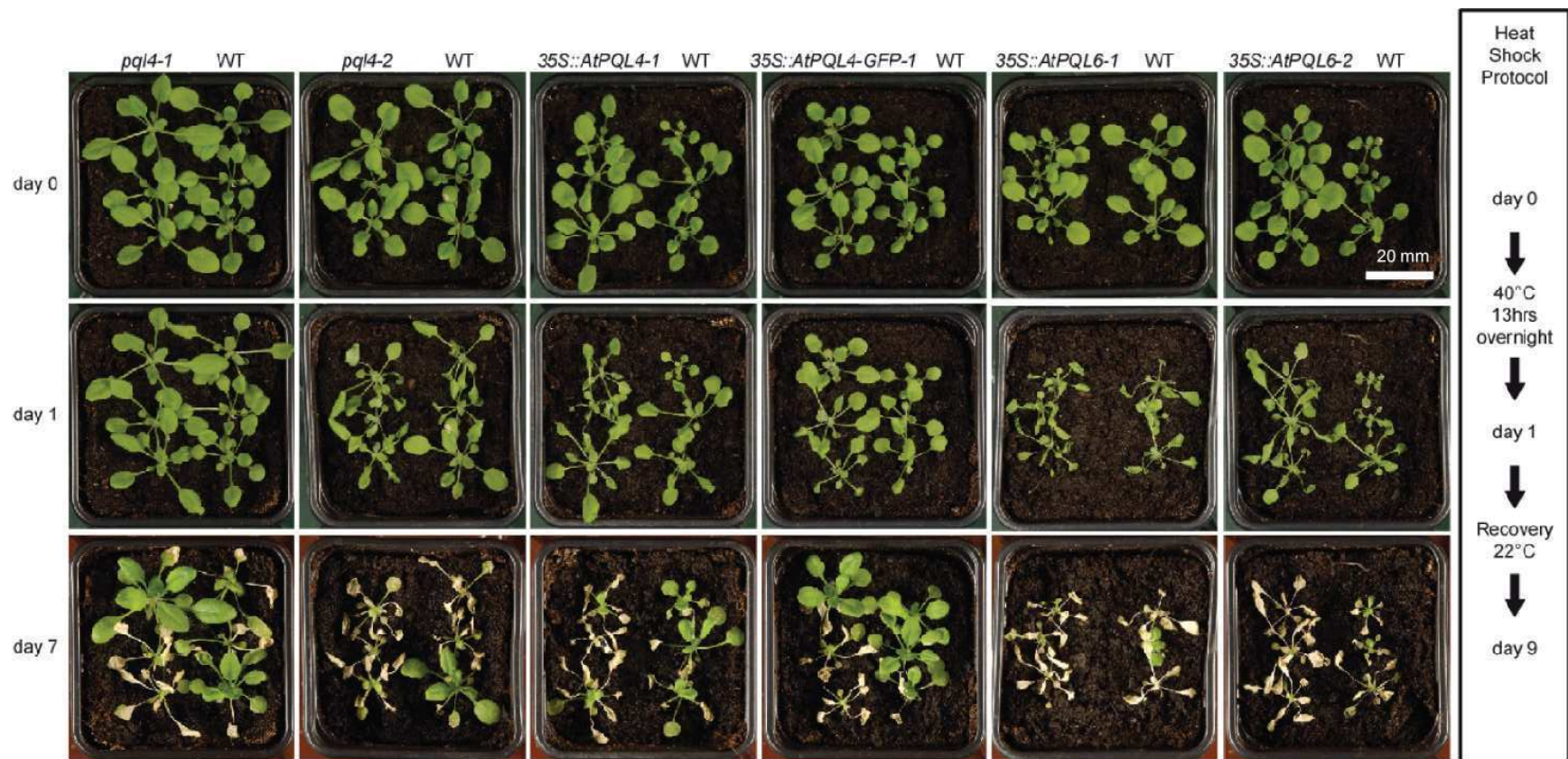


Figure 5-11. Heat shock of mutant lines with altered expression of *AtPQL4* and *AtPQL6*. Plants were grown in long-day conditions (16 hr light, 22 °C, 60% relative humidity / 8 hr dark, 18 °C, 70% relative humidity) for 3 weeks and then heat shocked at 40°C overnight before being allowed to recover at 21°C as described in the right hand-panel. Each pot contains three wild type plants and three plants from the indicated mutant line.

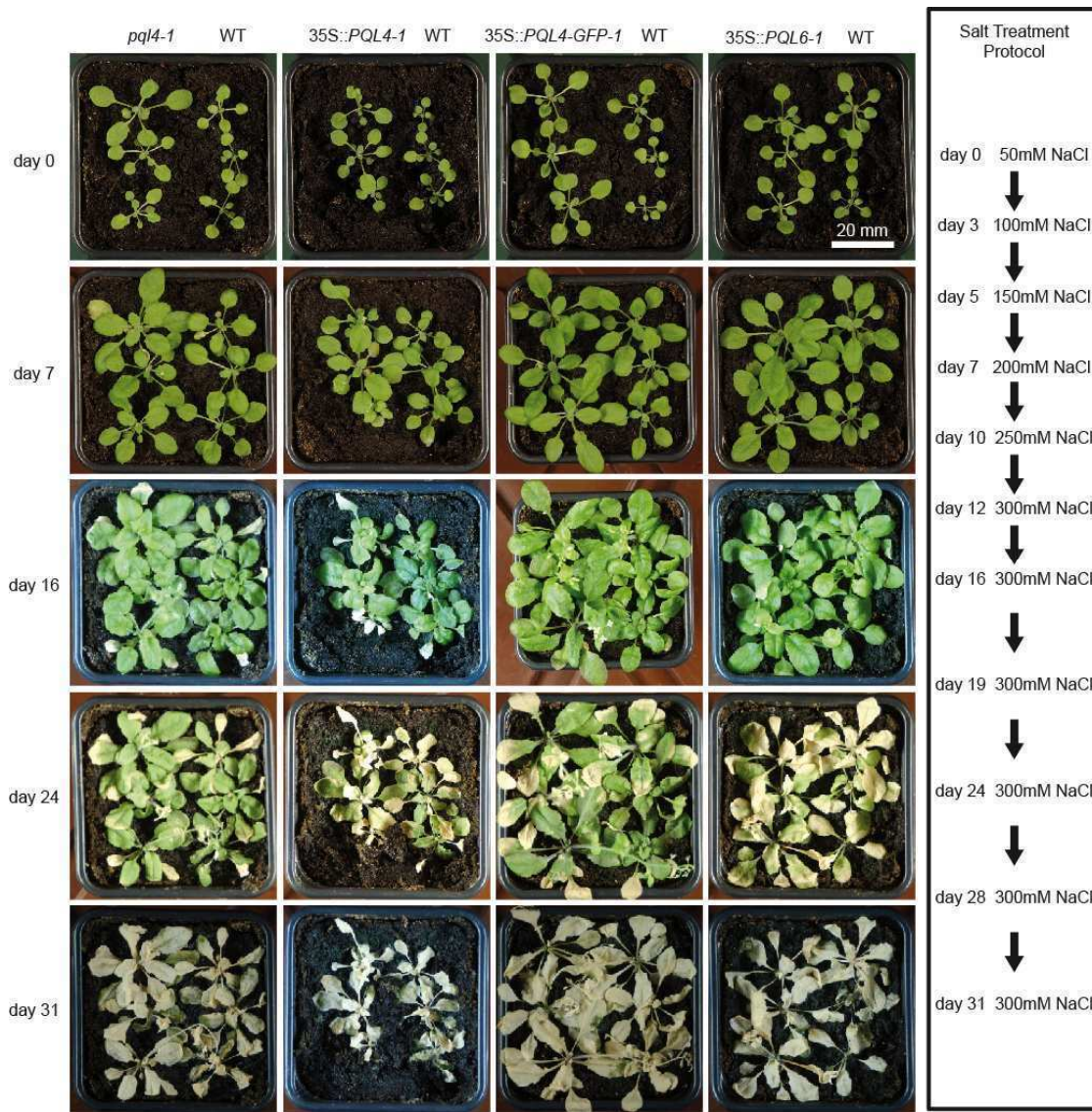


Figure 5-12. Salt stress treatment of mutant lines with altered expression of *AtPQL4* and *AtPQL6*. Plants were grown in long-day conditions (16 hr light, 22 °C, 60% relative humidity / 8 hr dark, 18 °C, 70% relative humidity) and watered regularly for 3 weeks. After 3 weeks they were watered with increasing concentrations of sodium chloride as indicated in the right-hand panel and photographed at regular intervals. Each pot contains three wild type and three plants of the indicated mutant line.

5.3.4.4 Sensitivity to salt: root-bending assay

To further investigate salt tolerance in one mutant, *pql4-1*, a root-bending assay was conducted. *AtPQL4* has a putative function in glycosylation. Plants were grown on “control medium” (see materials and methods) for 4 days before being transferred to new plates containing control medium or media supplemented with 100 or 200 mM NaCl. The

plates were then turned upside-down and the plants allowed to grow for a further 3 days before being photographed (see figure 5-13). Root growth on the new medium can be assessed by comparing the final position of the root tip (yellow arrow) with the position of the root tip at reorientation (red arrow). Root growth after transfer was greatest on control medium. Root growth on media containing 100 mM NaCl was approximately one-third that on non-supplemented media and at 200 mM NaCl root growth was almost completely inhibited. Again however, no differences could be seen between wild-type and *pql4-1* plants.

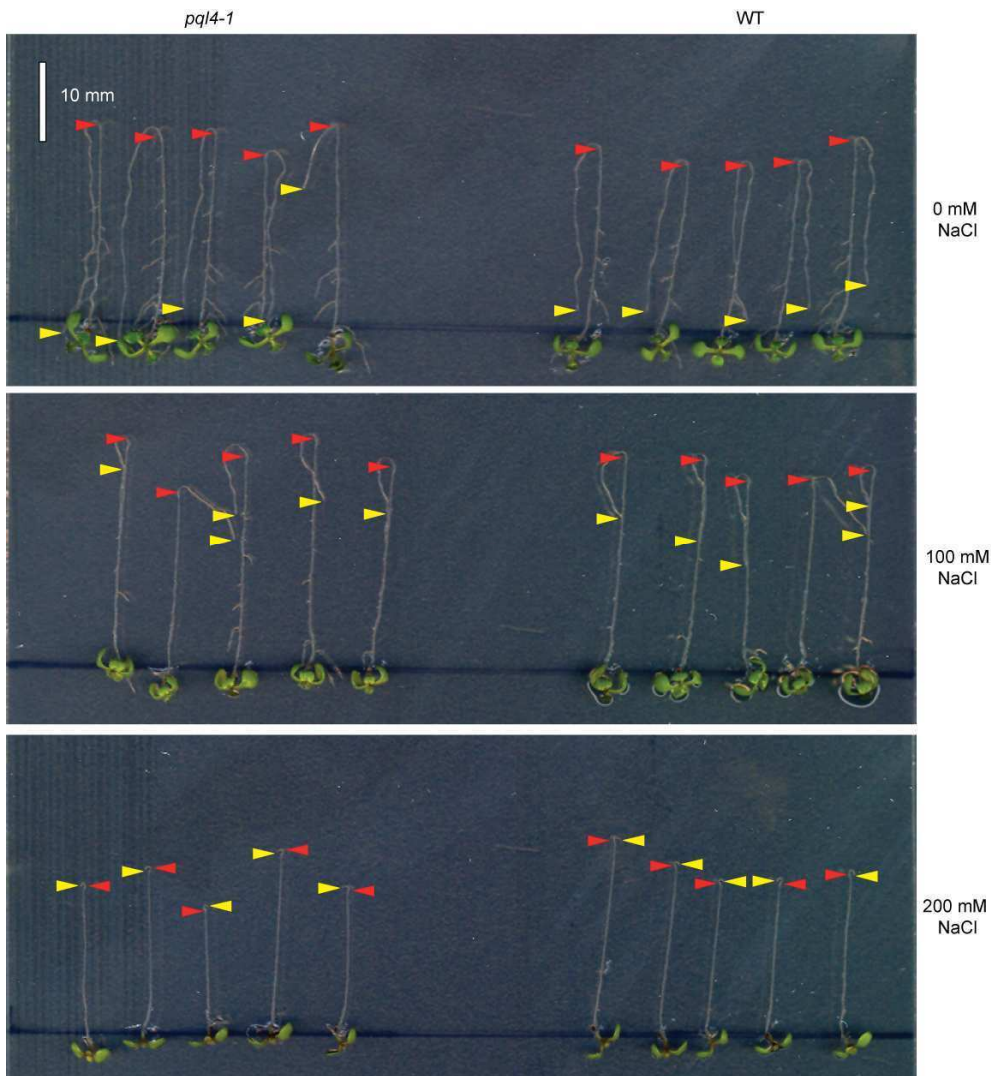


Figure 5-13. Root-bending assay of *pql4-1* salt tolerance. Wild-type and *pql4-1* seedlings were grown on control medium based on Arteca and Arteca (2000) in long day conditions (16 hr light/8 hr dark, 21 °C) for four days before being transferred to new plates containing 0, 100 or 200 mM NaCl and turned upside-down. Red arrows indicate root position at re-orientation and yellow arrows the final position four days later.

5.3.4.5 Sensitivity to salt: seedling salt and osmotic stress tolerance assay

For further investigation of salt and osmotic tolerance, some mutant lines were grown alongside wild type on media designed to inflict salt or osmotic stress at the point of germination. Unlike the root-bending assay described above, where the stress was applied on 4-day-old seedlings, here the seeds were sown directly on media supplemented with 50 mM NaCl or 50 mM KCl.

The three mutants tested were *pql3-1*, *pql4-1* and *pql5-1*. The results of this experiment are shown in figure 5-14. Noticeable differences were seen between media. After 14 days the roots of plants on control media had reached the bottom of the plate and several lateral roots were visible. On media supplemented with 50 mM KCl overall root length was similar but there were less lateral roots. Roots of plants on media supplemented with 50 mM NaCl were much smaller and stopped growing after 4-5 days. Shoots were smaller on both 50 mM NaCl and 50 mM KCl compared to control media.

Again however, no differences were observed between wild type and any of the mutants tested. The percentage water content, as determined by fresh and dry weight measurements, was similar in each mutant line (figure 5-14b).

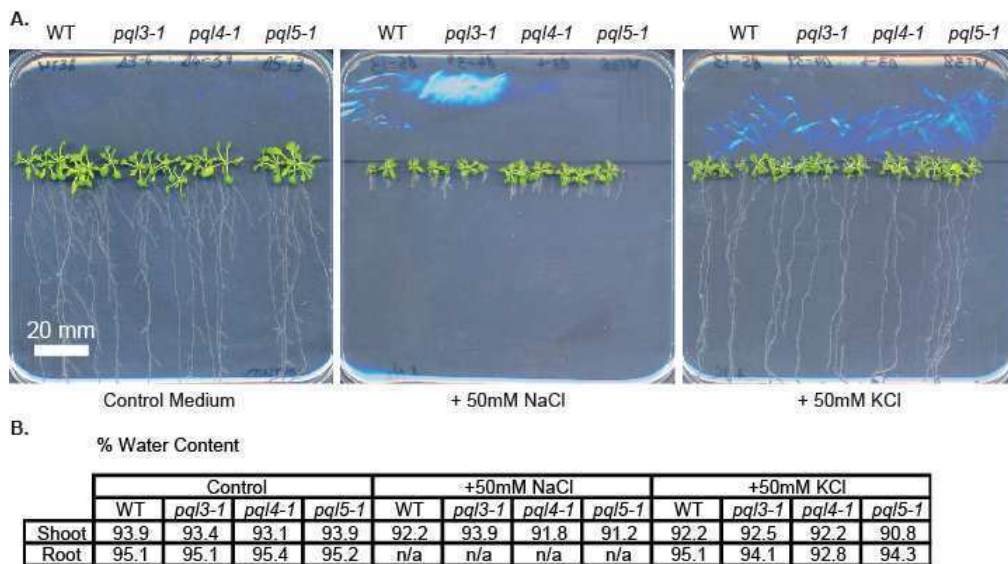


Figure 5-14. Seedling salt and osmotic stress tolerance assay. Wild type and the indicated mutant lines were grown on control medium based on Arteca and Arteca (2000) in long day conditions (16 hr light/8 hr dark, 21 °C) or the same media supplemented with 50mM NaCl or 50mM KCl. **A.** Images taken 14 days after germination. **B.** Percentage water content of roots and shoots as determined by fresh and dry weight measurements.

5.3.4.6 Sensitivity to dehydration

To check for altered sensitivity to dehydration each line was grown on soil and then allowed to dry out. Plants were watered regularly for three weeks, before watering was stopped. Example images of *AtPQL4* and *AtPQL6* mutant lines are shown in figure 5-15. After two weeks, a noticeable dehydration effect was apparent. Leaves appeared thinner and lacked turgor and several plants had started to flower. After three weeks, no further growth occurred and the plants could not be recovered by resuming watering. No difference in symptoms was observed between wild type and any of the mutant lines.

5.3.4.7 Analysis of leaf water loss

For one mutant line, *pql5-1*, dehydration was investigated further by examining water loss from excised leaves. *AtPQL5* is a tonoplast protein with relatively strong expression in rosette leaves (see chapter 4). Rosette leaves were excised from each line and allowed to dehydrate under controlled conditions (growth room at 21 °C and 60 % relative humidity). As can be seen in figure 5-16 both wild-type and *pql5-1* leaves lost almost half their initial fresh weight over the course of the experiment (460 minutes). However, no differences were seen in the rate of water loss.

5.3.4.8 Sensitivity to cadmium

To test cadmium tolerance, mutants were grown on media containing 0, 25 or 50 μM CdCl_2 . To check for a possible influence of sulfur nutrition, the experiments were carried out on both control (contains just over 0.5 mM SO_4^{2-}) and sulfur-deplete medium (1.96 μM SO_4^{2-}).

CdCl_2 had a detrimental effect on every line tested (figure 5-17). Compared to control medium, plants on media containing 25 μM CdCl_2 had smaller shoots, shorter roots and chlorotic leaves. At 50 μM CdCl_2 , the toxic effects were stronger, particularly when comparing root length. However, no difference in the severity of symptoms was seen between control and low sulfur medium. In addition, none of the mutant lines appeared more or less sensitive than wild type.

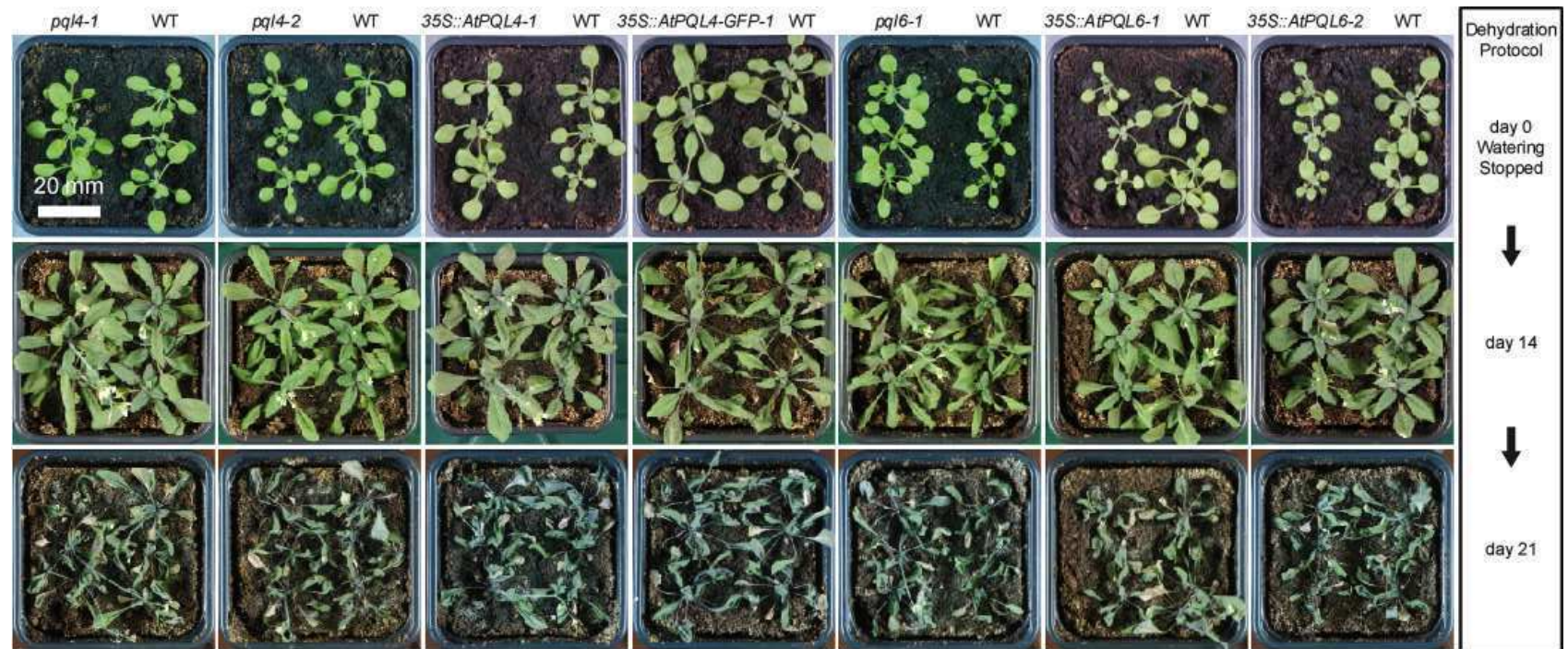


Figure 5-15. Dehydration of mutant lines with altered expression of *AtPQL4* and *AtPQL6*. Plants were grown in long-day conditions (16 hr light, 22 °C, 60% relative humidity / 8 hr dark, 18 °C, 70% relative humidity) and watered regularly for three weeks. After three weeks watering was stopped and the plants were photographed as indicated in the right-hand panel. Each pot contains three wild type and three plants of the indicated mutant line.

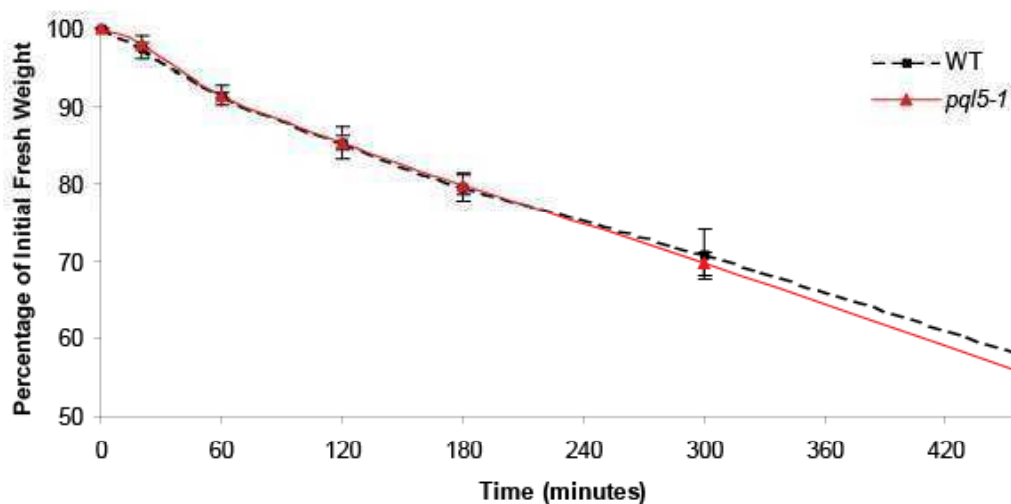


Figure 5-16. Water loss from excised leaves of wild-type and *pql5-1*. Wild type and *pql5-1* plants were grown in long-day conditions (16 hr light, 22 °C, 60% relative humidity / 8 hr dark, 18 °C, 70% relative humidity) and watered regularly for four weeks. After four weeks a medium sized rosette leaf was removed from three individual plants. The leaves were allowed to dry at 21°C in an unsealed petri dish and the fresh weight was measured at regular intervals. Data is presented as percentages of initial fresh weight. Dashed black lines represent wild type and red lines *pql5-1*. Error bars represent standard error.

5.3.4.9 Sensitivity to SNP

Sensitivity to the nitric oxide donor, SNP was tested on media containing 0, 25 or 75 μ M SNP. Again, the experiment was conducted on both control and low sulfur medium.

As shown in figure 5-18, a detrimental affect of SNP was found at 25 μ M. At 25 μ M SNP, roots appeared shorter than those on SNP-free media and shoots appeared smaller with narrower leaves. At 75 μ M SNP, although the seeds did germinate and emerging cotyledons were visible, growth stopped one to two days after germination. Again the severity of symptoms did not differ between control and low sulfur media and no differences were seen between any of the mutant lines and wild type.

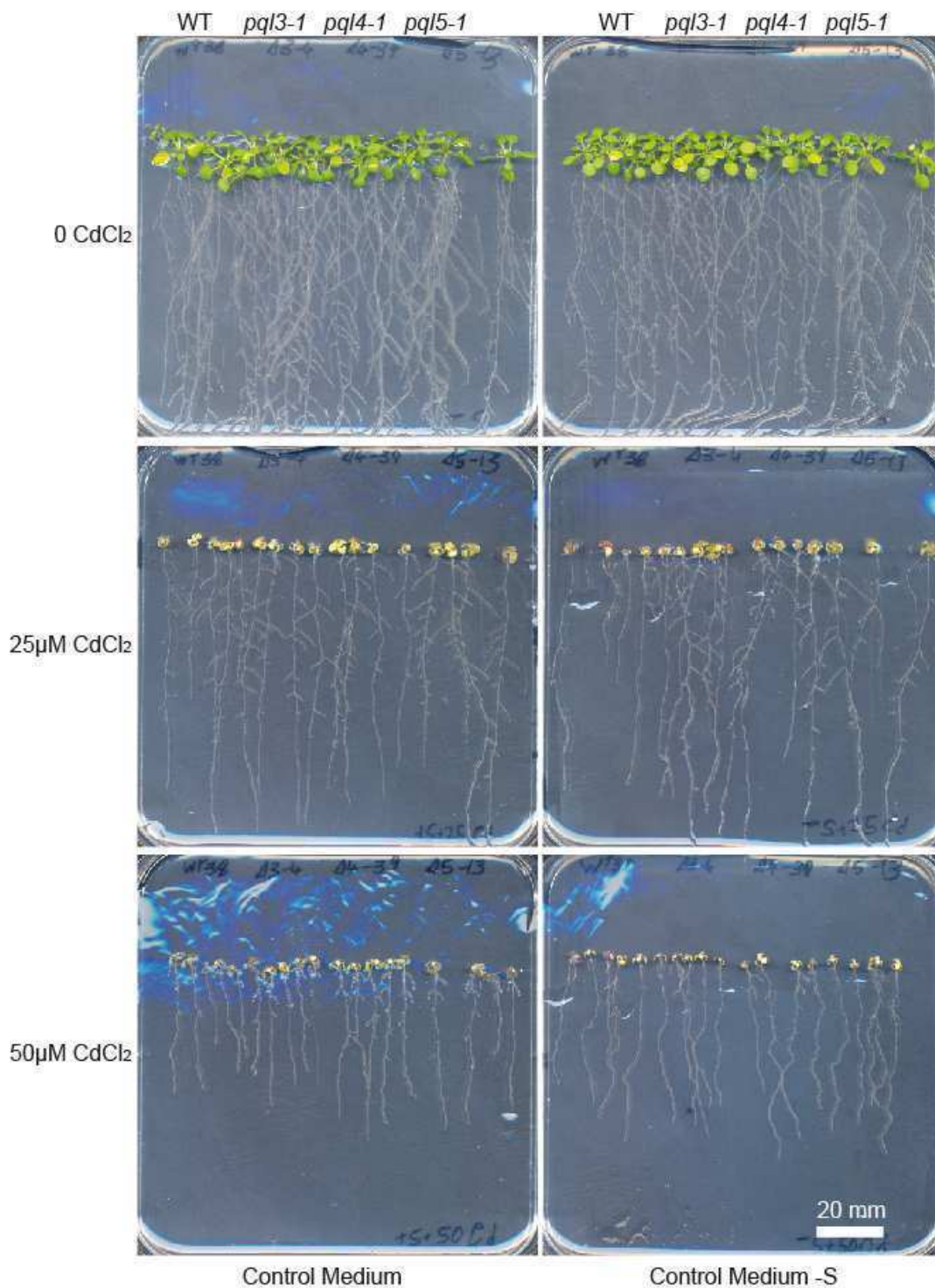


Figure 5-17. Cadmium tolerance of *AtPQL* mutants in control and low sulfur conditions. Wild type and five plants each of the indicated mutant lines were grown on agar plates in long day conditions (16 hr light/8 hr dark, 21 °C). Media was either “control media”, based on Arteca and Arteca (2000) containing just over 0.5 mM SO_4^{2-} , or “control media – S” where SO_4^{2-} was reduced to 1.96 μM . CdCl_2 was added at the indicated concentration. Photographs were taken 14 days after germination.

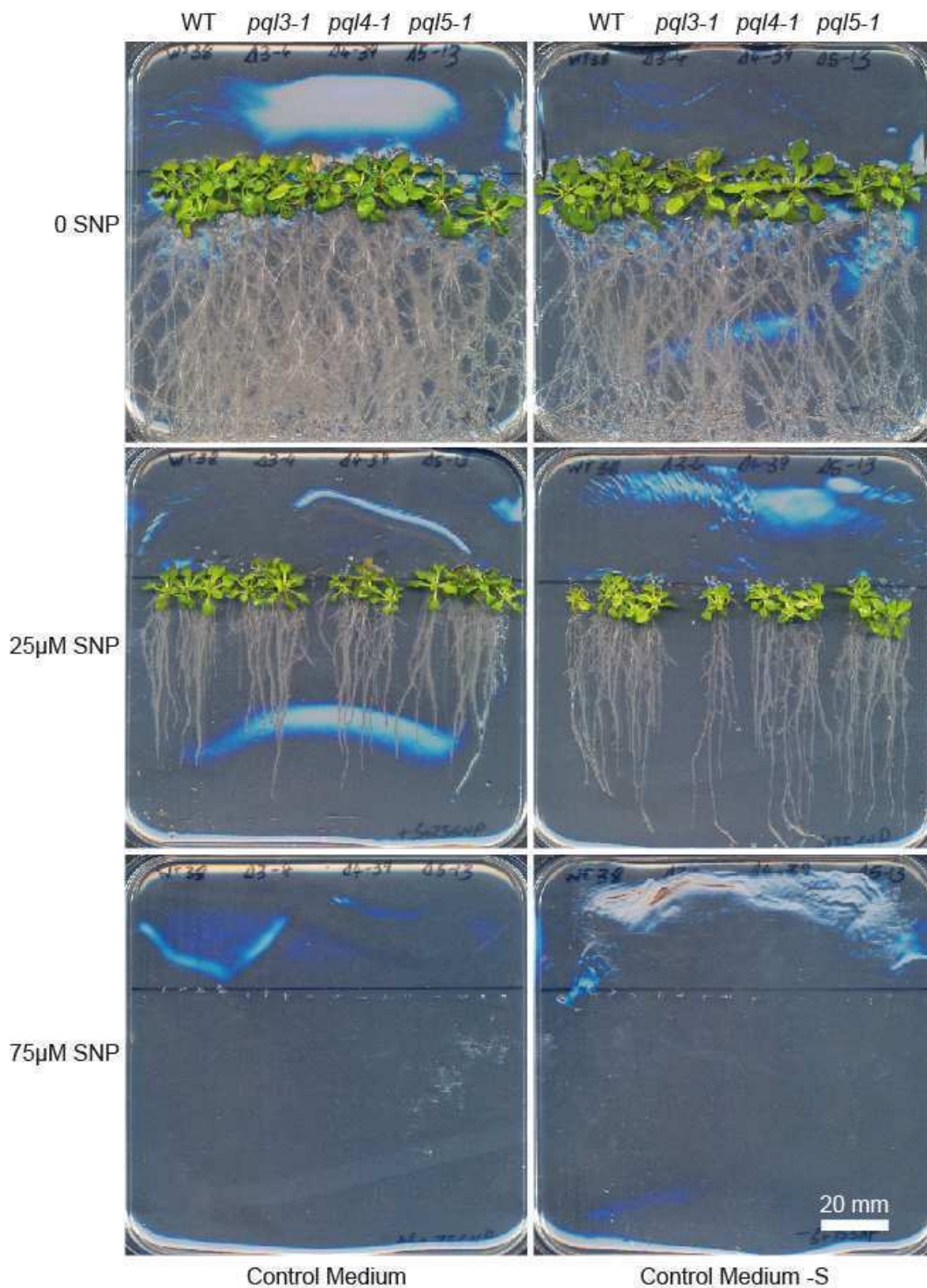


Figure 5-18. Effect of the nitric oxide donor SNP on *AtPQL* mutant lines in control and low sulfur conditions. Wild type and five plants each of the indicated mutant lines were grown on agar plates in long day conditions (16 hr light/8 hr dark, 21 °C). Media was either “control media”, based on Arteca and Arteca (2000) containing just over 0.5 mM SO_4^{2-} , or “control media – S” where SO_4^{2-} was reduced to 1.96 μM. SNP was added at the indicated concentration. Photographs were taken 21 days after germination.

5.3.4.10 Analysis of pollen viability and germination

For each *AtPQL* mutant examined, homozygotes could be isolated indicating that none were gamete lethal. However, it was still possible that the mutants had a gamete-specific phenotype. For *pql4-1*, *pql4-2* and *pql6-1* the progeny of heterozygous parent plants were genotyped by PCR to check for altered segregation of the knockout allele. A “normal” polymorphism, not affecting transmission of genetic material to the next generation, segregates in a 1:2:1 ratio (homozygous to heterozygous to wild type). Distortions from this ratio indicate altered transmission. In the progeny of both *pql4* heterozygotes, there were less homozygous and more wild type plants than expected (wild type to heterozygous to homozygous ratio was 9:17:4 for *pql4-1* and 9:14:3 for *pql4-2*). χ^2 tests for each individual line do not indicate a significant difference from the expected 1:2:1 ratio. However, if the data for both *pql4* lines is treated cumulatively (18:31:7) then a p-value of 0.08 is obtained.

To check for viability, pollen was stained with fluorescein diacetate (FDA) and viewed using confocal microscopy. Viable and non-viable pollen were distinguishable by the presence or absence of fluorescence (see figure 5-19). For each line tested, almost all pollen fluoresced after incubation for 30 minutes with FDA indicating high levels of pollen viability. None of the mutants had noticeably lower viability than wild type.

As a further test of pollen quality, *in vitro* germination assays were carried out for some lines. Pollen was isolated from individual flowers and applied to pollen germination medium. As reported by other laboratories, germination rates varied considerably from day to day (Johnson-Brousseau and McCormick, 2004). In order to compare mutant and wild type, care was taken to use plants and flowers of the same age and all replicates were performed on the same day. As shown in figure 5-20a, after optimisation, pollen germination rates of between 60 and 80% could be achieved. However, none of the mutant lines tested showed significantly different germination rates than wild type. Images of germinated pollen are shown in figure 5-20b. No differences in morphology could be found in either pollen grains or tubes, which appeared to be the same shape and length in wild type and all the mutants tested.

Table 5-3. Zygosity testing of the progeny of heterozygous *pql4* and *pql6* mutants. A parent plant heterozygous for the *pql4-1*, *pql4-2* or *pql6-1* allele was allowed to self-fertilise and the number of progeny plants homozygous, heterozygous or wild type with respect to the transgene is indicated. The row *PQL4/pql4* (both) is the combined data for the two individual *pql4* insertion lines. A χ^2 test was used to test for significant differences from the 1:2:1 ratio expected for a normal polymorphism not affecting genetic transmission. The p-value indicates the likelihood that the observed ratios could occur by chance if the mutation does not affect transmission.

Parent Plant	Progeny of Heterozygote			χ^2 test p-value normal (1:2:1)
	wild type	heterozygous	homozygous	
<i>PQL4/pql4-1</i>	9	17	4	0.333
<i>PQL4/pql4-2</i>	9	14	3	0.232
<i>PQL4/pql4</i> (total)	18	31	7	0.084
<i>PQL6/pql6-1</i>	8	13	5	0.707

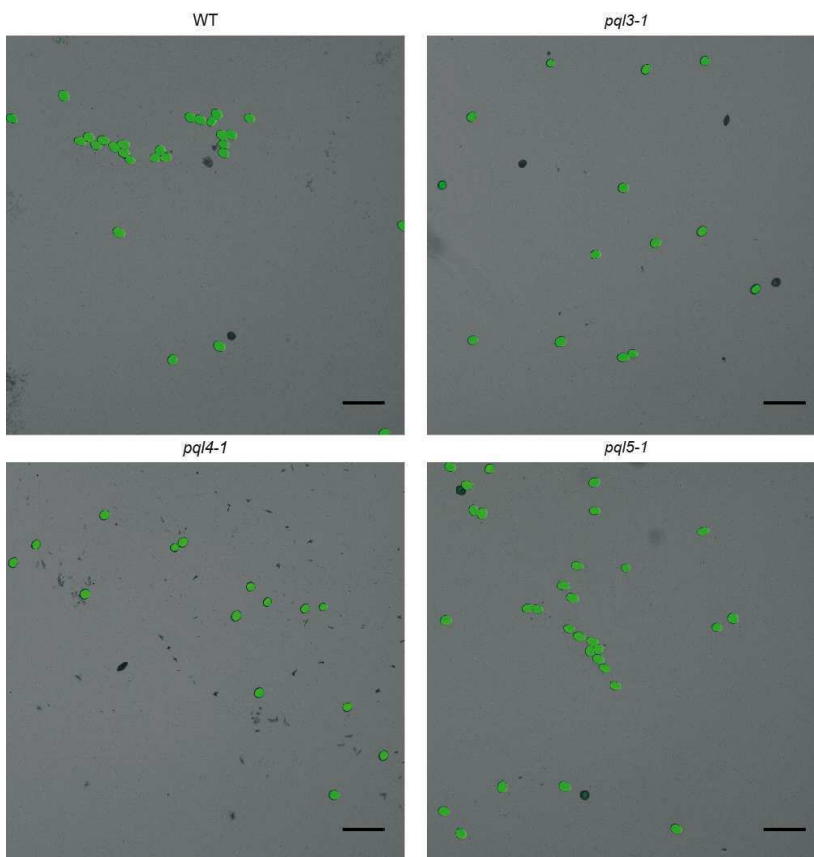


Figure 5-19. Viability of *AtPQL* mutant pollen. Pollen was harvested from six-week old wild-type or *AtPQL* mutant plants and stained with fluorescein diacetate (FDA). Recently opened flowers were harvested and pollen isolated by dabbing anthers onto a microscope slide containing pollen germination medium. Pollen was stained with 0.1mg/ml FDA for 30 minutes and viewed using confocal microscopy with standard settings for FITC. Viable and inviable pollen can be distinguished by the presence or absence of fluorescence. Scale bars represent 100 μ m.

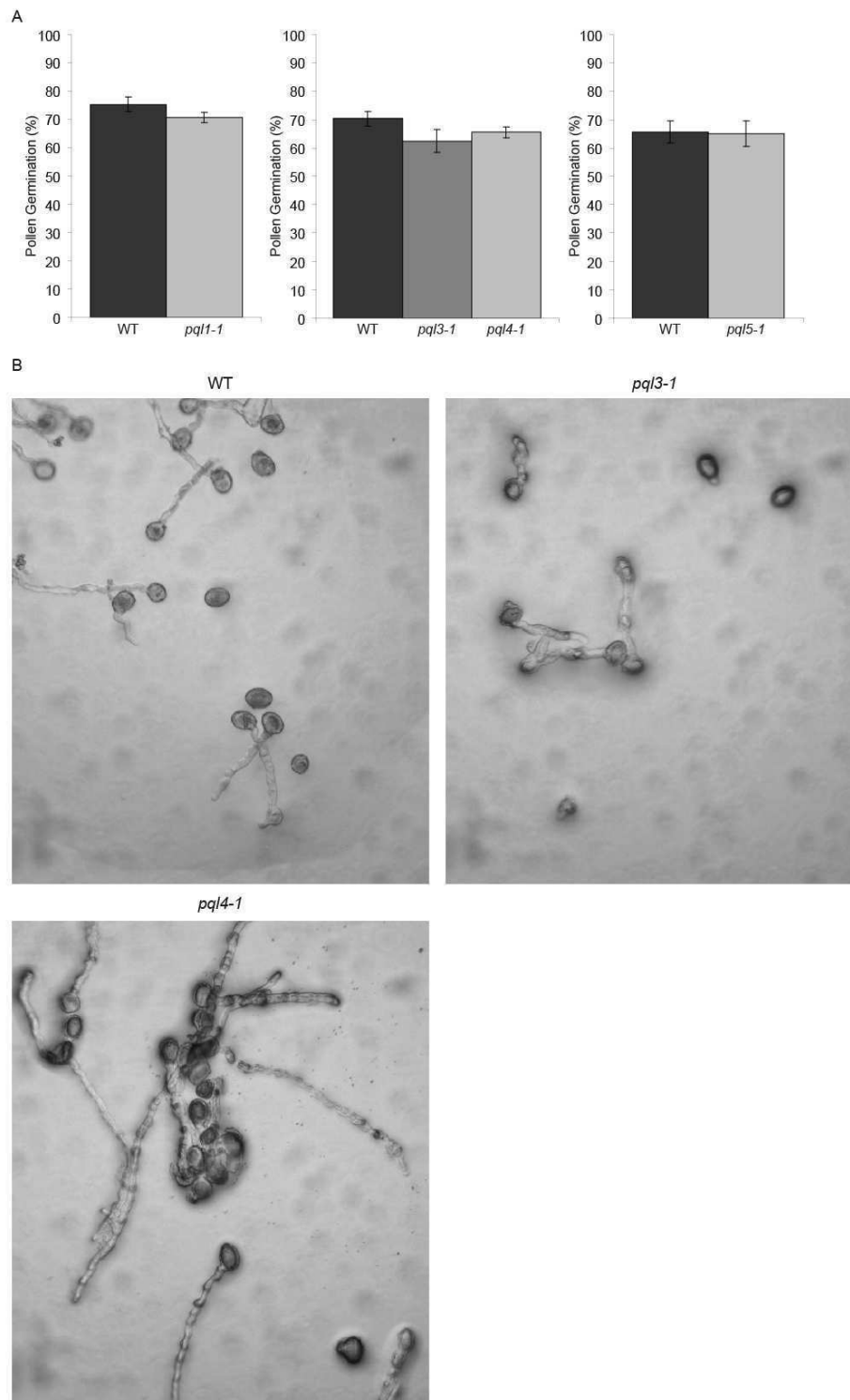


Figure 5-20. *In vitro* germination of *AtPQL* mutant pollen. Pollen was harvested from six-week old wild-type or *AtPQL* mutant plants. Recently opened flowers were harvested and pollen isolated by dabbing anthers onto a petri dish containing pollen germination medium. **A.** Percentage *in vitro* germination of pollen harvested from wild-type and *AtPQL* mutant flowers. The proportion of germinated pollen was counted after incubation overnight. Error bars represent standard errors ($n=3$ for *pq1-1*, $n=10$ for *pq3-1* and *pq4-1*, $n=5$ for *pq5-1*) **B.** Images of pollen after overnight incubation on pollen germination medium.

5.4 Results: Transcriptional profiling of mutant lines

5.4.1 Microarray experiments

Microarrays were used to profile gene expression in four different mutants - *pql4-1*, *pql4-2*, *pql6-1* and *35S::AtPQL6-1*. *pql4-1* and *pql4-2*, contain a T-DNA insertion in *AtPQL4*. *pql6-1* contains a T-DNA insertion in an intron of *AtPQL6*. Although *AtPQL6* transcript is still produced it is less than half as abundant as in wild type (see figure 5-7b). Finally, *35S::AtPQL6-1* is an over-expresser mutant which produces approximately seven times as much transcript as wild type (see figure 5-9c). *AtPQL4* and *AtPQL6* are closely related to LEC35/MPDU1 and are predicted to be involved in Man/Glc-P-Dol dependent glycosylation.

To compare transcriptional profiles, RNA samples from 12-day old seedlings grown under control conditions (see materials and methods) were analyzed. Three biological replicates were performed for each mutant line and wild type. For each mutant, data was analysed by rank product (RP) and iterative group analysis (iGA). A list of genes sorted by RP is presented for each mutant in appendix B. All genes with a false detection rate (FDR) of 10% or less were used for further analyses.

Table 5-4 shows the number of genes significantly up- or down-regulated in each mutant relative to wild type. The numbers varied considerably between mutant lines. Relatively few differentially regulated genes were seen in the *pql4* knockout mutants (two in *pql4-1* and 40 in *pql4-2*) compared to the *AtPQL6* mutants (89 in *pql6-1* and 381 in *35S::AtPQL6-1*).

The fact that so few differentially regulated genes were identified in *pql4-1* is likely due to the presence of one outlier amongst the three replicates (see appendix B). However, despite the fact that few individual genes showed significant differential regulation between wild type and *pql4-1*, iGA did show significant changes in the expression of groups of functionally related genes.

Table 5-4. Number of differentially regulated genes in *AtPQL* mutants. Genes with differential regulation in *pql4-1*, *pql4-2*, *pql6-1* and *35S::AtPQL6-1* mutant plants compared to wild type were identified using the Affymetrix ATH1 microarray to profile transcript level in 12-day-old seedlings. Plant material was grown on agar plates containing control medium based on Arteca and Arteca (2000) in long day conditions (16 hr light/8 hr dark, 21 °C). Three replicate chips were hybridised for each line and analysed using the rank product method. The total number of genes up- or down-regulated relative to wild type with a false detection rate of below 1% or 10% is given.

FDR (%)	Line							
	<i>pql4-1</i>		<i>pql4-2</i>		<i>pql6-1</i>		<i>35S::AtPQL6-1</i>	
	down	up	down	up	down	up	down	up
<1	0	0	5	0	1	19	2	39
<10	0	2	32	8	7	82	21	360

5.4.2 Altered gene expression in *AtPQL4* and *AtPQL6* mutants

The number of genes showing common or opposite regulation amongst mutants is shown in figure 5-21. Given that *pql4-1* and *pql4-2* are different knockout alleles of the same gene, several common changes could be expected. In contrast, several opposite changes (i.e. genes up-regulated in one but down-regulated in the other) might be expected in *pql6-1* and *35S::AtPQL6-1*, given that one is an *AtPQL6* knockout and the other an over-expresser.

Of the two genes up-regulated in *pql4-1*, one was also up-regulated in *pql4-2*. This gene, an uncharacterised lipid transfer like protein, also showed increased expression in *35S::AtPQL6-1*.

Only two genes showed opposite regulation in *pql6-1* and *35S::AtPQL6-1* (i.e. increased expression in one, decreased expression in the other). One was *AtPQL6* itself which was the most strongly down-regulated gene in *pql6-1* and the most strongly up-regulated gene in *35S::AtPQL6-1*. The other was *XTR7*, encoding a putative xyloglucan endotransglycosylase, which showed increased transcript level in *pql6-1* and decreased transcript level in *35S::AtPQL6-1*.

A greater number of genes showed common regulation between *pql6-1* and *35S::AtPQL6-1*. In total, eleven genes were up-regulated in both. Amongst these genes were two *HSP70* family members, one of which, *BiP3* encodes an ER-localised protein folding chaperone (Noh et al., 2003). A number of defence-related genes, including *PR-1*

and *HR4*, were also up-regulated in both mutants. Given the central role of *BiP3* in ER stress (see section 5.5.3.2) the up-regulation of this gene was confirmed by quantitative PCR (see figure 5-22).

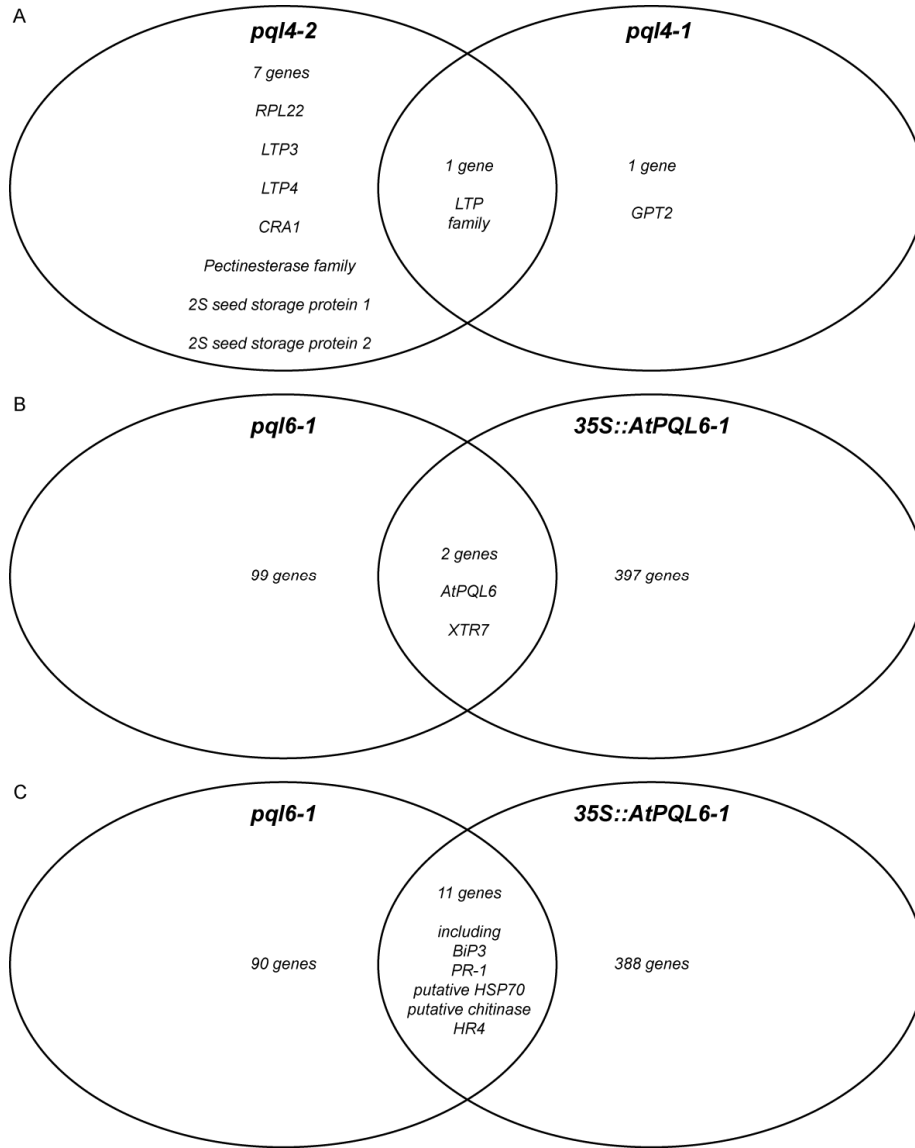


Figure 5-21. Genes showing common or opposite regulation in *AtPQL* mutants. Venn diagrams indicate the number of genes with differential regulation relative to wild type. Venns are based on rank product lists of up and down regulated genes in microarray experiments comparing gene regulation in mutant and wild type (see table 5-4). Differentially regulated genes with a false detection rate of less than 10% are indicated. **A.** Genes up-regulated in *pql4-1*, *pql4-2* or both (indicated in overlap). **B.** Differentially regulated genes in *pql6-1* or *35S::AtPQL6-1*. Genes with opposite responses (up in one down in the other) are indicated in the overlap. **C.** Differentially regulated genes in *pql6-1* or *35S::AtPQL6-1*. Genes with similar responses (i.e. up in both or down in both) are indicated in the overlap.

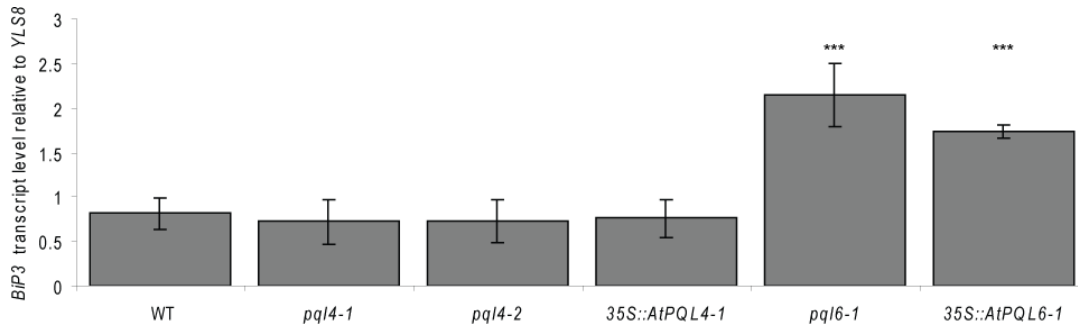


Figure 5-22. Differential regulation of *BiP3* in *AtPQL4* and *AtPQL6* mutants. *BiP3* transcript level relative to *YLS8* was determined by real-time reverse-transcription PCR in 12-day old seedlings grown on agar plates containing control medium based on Arteca and Arteca (2000) in long day conditions (16 hr light/8 hr dark, 21 °C). The average value of three biological replicates is given, error bars represent standard error. A t-test was used to compare each mutant line and wild type. *** = p-value < 0.001.

5.4.3 Iterative group analysis (iGA)

One of the most noticeable differences between the mutants and wild type was altered expression of genes encoding lipid transfer proteins (LTP) and LTP homologues. *LTP* genes were prominent amongst the list of genes with altered expression yet few individual family members showed differential regulation in more than one mutant. Therefore, such common regulation of function in the mutants (e.g. alteration of lipid transfer protein activity) would not be apparent in Venn diagrams and other gene-specific analyses. To assess regulation of groups of functionally related genes, iterative group analysis (iGA) was used (see section 5.4.3).

iGA revealed several groups of functionally related genes with altered expression in each mutant. These groups are listed in tables 5-5 and 5-6. One group related to lipid transfer (which consists of *LTP* and *LTP* related genes) was up-regulated in all three knockout mutants and down-regulated in the over-expresser. In addition, a group of wounding-related genes showed decreased expression in both *pql4-1* and *pql4-2*. Several groups showed common regulation in *pql6-1* and one of the *pql4* knockout alleles. These included transcripts for defence-related transport-associated proteins (down in *pql4-2* and *pql6-1*), toxin-binding proteins (down in *pql6-1* and *pql4-1*), TATA box-binding association factors (down in *pql6-1* and *pql4-1*) and plant defensins (up in both *pql4-2* and *pql6-1*).

Table 5-5. Groups of functionally related genes showing differential regulation in *AtPQL4* mutants. Groups of functionally related genes showing differential regulation in *AtPQL4* mutants were identified by iterative group analysis (iGA) based on biased distribution of group members in rank product lists of up and down regulated genes in microarray experiments comparing gene regulation in mutant and wild type (see table 5-4). A minimum p-value is given for each group representing the statistical confidence with which the group can be identified by iGA. Groups associated with defence and cell surface modification are highlighted orange and groups associated with lipid biosynthesis and metabolism are highlighted green.

A. Up in *pql4-1*

Top Changed Groups	P-value
Response to copper ion (GO 0046688)	6.50E-06
Lipid transport (GO 0006869)	3.20E-05
nitrogen transport-associated proteins (MapMan)	1.20E-04
copper/zinc superoxide dismutase activity (GO 0004785)	1.40E-04
senescence (GO 0010149)	1.80E-04

B. Down in *pql4-1*

Top Changed Groups	P-value
toxin receptor binding (GO 0050827)	6.80E-07
flavonoid biosynthetic process (GO 0009813)	1.30E-06
response to wounding (GO 0009611)	7.10E-05
TATA box binding protein associated factor (IPR 004823)	1.40E-04

C. Up in *pql4-2*

Top Changed Groups	P-value
Lipid transport (GO 0006869)	7.70E-12
nutrient reservoir activity (GO 0045735)	1.20E-09
Seed oilbody biogenesis (GO 0010344)	1.30E-04
Plant defensins superfamily (TAIR GeneFamily)	1.40E-04

D. Down in *pql4-2*

Top Changed Groups	P-value
Response to jasmonic acid stimulus (GO 0009753)	1.20E-06
response to wounding (GO 0009611)	1.70E-05
squalene epoxidase-like family (TAIR GeneFamily)	5.10E-05
Pathogen response and transport associated proteins (MapMan)	9.30E-05
acyl lipid metabolism (predicted mitochondrial) (TAIR GeneFamily)	1.10E-04
C-22 sterol desaturase activity (GO 0000249)	1.20E-04
proton-transporting ATP synthase complex coupling factor F(o) (GO 0045263)	1.80E-04

Table 5-6. Groups of functionally related genes showing differential regulation in *AtPQL6* mutants. Groups of functionally related genes showing differential regulation in *AtPQL6* mutants were identified by iterative group analysis (iGA) based on biased distribution of group members in rank product lists of up and down regulated genes in microarray experiments comparing gene regulation in mutant and wild type (see table 5-4). A minimum p-value is given for each group representing the statistical confidence with which the group can be identified by iGA. Groups associated with defence and cell surface modification are highlighted orange and groups associated with lipid biosynthesis and metabolism are highlighted green.

A. Up in *pql6-1*

Top Changed Groups	P-value
defense response to fungus (GO 0050832)	3.10E-11
systemic acquired resistance (MapMan)	3.20E-06
Plant defensins superfamily (TAIR GeneFamily)	1.10E-05
response to biotic stimulus (GO 0009607)	1.20E-05
antiporter activity (GO 0015297)	1.50E-05
lipid transport (GO 0006869)	2.00E-05
xenobiotic metabolic process (GO 0006805)	2.60E-05
chitin catabolic process (GO 0006032)	5.20E-05
WRKY transcription factor family IIA (TAIR GeneFamily)	8.30E-05
COBRA protein family (TAIR GeneFamily)	1.30E-04

B. Down in *pql6-1*

Top Changed Groups	P-value
nucleosome (GO 0000786)	8.60E-06
toxin receptor binding (GO 0050827)	4.50E-05
TATA box binding protein associated factor (TAF) (IPR004823)	1.30E-04
indole-3-acetic acid amido synthetase (GO 0010279)	1.30E-04
Pathogen response and transport associated proteins (MapMan)	1.60E-04
histone phosphorylation (GO 0016572)	1.60E-04

C. Up in *35S::AtPQL6-1*

Top Changed Groups	P-value
cohesin complex (GO 0008278)	1.30E-04
peptidyl-prolyl cis-trans isomerase (TAIR GeneFamily)	1.30E-04
DNA topoisomerase type I activity (GO 0003917)	1.60E-04
copper/zinc superoxide dismutase activity (GO 0004785)	1.60E-04
isopentenyl diphosphate biosynthetic process (GO 0019288)	1.70E-04

D. Down in *35S::AtPQL6-1*

Top Changed Groups	P-value
acyl lipid metabolism (Ketoacyl-CoA Synthase) (TAIR GeneFamily)	3.80E-07
lipid transport (GO 0006869)	1.00E-06
structural constituent of cell wall (GO 0005199)	8.50E-05
GPI anchor binding (0048503)	1.10E-04
UDP-glucose 6-dehydrogenase activity (GO 0003979)	1.10E-04
C-22 sterol desaturase activity (GO 0000249)	1.50E-04
gametophyte development (GO 0048229)	1.70E-04
Monosaccharide transporter-like gene family (TAIR GeneFamily)	1.80E-04

5.4.4 Further analysis of differentially regulated genes

Given that *AtPQL4* and *AtPQL6* are proposed to function in glycosylation and that glycosylation is a common modification amongst proteins targeted to the secretory pathway, I analysed the sub-cellular localisation and N-glycosylation status of the proteins encoded by differentially regulated genes (i.e. those with an FDR below 10 %; see table 5-7).

Sub-cellular localisation: In total, 17% of nuclear-encoded genes in the *A. thaliana* genome are predicted to be targeted to the secretory pathway (see materials and methods). For all three knockouts, genes predicted to encode secretory pathway proteins featured more frequently amongst the list of differentially regulated genes than expected by chance (between 33 and 86% of each gene list; see table 5-7a). In *AtPQL6* over-expressing mutants, genes predicted to encode secretory pathway associated proteins were over-represented (72 %) amongst down-regulated genes whilst being under-represented (8 %) amongst up-regulated genes. A chi-squared test confirmed that these differences could not be explained by random variation (see colour coding in table 5-7).

N-glycosylation status: In total 65 (12.8 %) of the 509 differentially regulated genes in the *AtPQL* mutants are predicted to encode potential N-glycoproteins (see table 5-7b). Using the same criteria approximately 16 % of all *A. thaliana* genes are predicted to encode N-glycoproteins (based on a random sample of over 200 protein sequences – see materials and methods).

However, the proportion of genes predicted to encode N-glycoproteins amongst all differentially regulated genes varied considerably between mutants. Strikingly, of the genes which are differentially regulated in *pql6-1* and down-regulated in *35S::AtPQL6-1*, significantly more than expected encode predicted N-glycoproteins (between 50 and 67.1 %). Conversely, of the up-regulated genes in *35S::AtPQL6-1*, significantly less than expected encode predicted N-glycoproteins (4.7%).

Table 5-7. Predicted sub-cellular localisation (**A**) and N-glycosylation status (**B**) of proteins encoded by differentially regulated genes in *AtPQL4* and *AtPQL6* mutants. Analysis is based on rank product lists of up and down regulated genes in microarray experiments comparing gene regulation in mutant and wild type (see table 5-4). Differentially regulated genes with a false detection rate of less than 10% are included. Given are the numbers of differentially regulated genes within each category (e.g. targeted to the chloroplast) and their percentage of the total number of differentially regulated genes in each line (FDR < 10%). The predicted sub-cellular localisation and N-glycosylation status was determined using TargetP and NetNGlyc (see materials and methods for details). In each case a χ^2 test was conducted to assess if each category was significantly over or under-represented compared to expected values (see text). Significantly over-represented categories (p-value < 0.05) are coloured red and significantly under-represented categories are coloured green.

A	<i>pql4-1</i>		<i>pql4-2</i>		<i>pql6-1</i>		<i>35S::AtPQL6-1</i>		expected
	down	up	down	up	down	Up	down	up	
Chloroplast	0	1 (50%)	3 (10%)	0	0	4 (4.9%)	1 (5.6%)	36 (10.5%)	8.5%
Mitochondrion	0	0	0	0	0	1 (1.2%)	0	5 (1.5%)	1.4%
Secretory Pathway	0	1 (50%)	10 (33.3%)	6 (85.7%)	3 (50%)	39 (47.6%)	13 (72.2%)	28 (8.2%)	17.4%
Undefined	0	0	11 (36.7%)	1 (14.3%)	2 (33.3%)	27 (32.9%)	3 (16.7%)	175 (51.2%)	49.9%
Other (cytoplasm etc.)	0	0	6 (20%)	0	1 (16.7%)	11 (13.4%)	1 (5.6%)	98 (28.7%)	22.8%
Total	0 (100%)	2 (100%)	30 (100%)	7 (100%)	6 (100%)	82 (100%)	18 (100%)	342 (100%)	100%

B	<i>pql4-1</i>		<i>pql4-2</i>		<i>pql6-1</i>		<i>35S::AtPQL6-1</i>		expected
	down	up	down	up	down	Up	down	up	
Predicted N-glycoproteins	0	0	6 (20%)	3 (42.9%)	3 (50%)	27 (32.9%)	9 (50%)	17 (5%)	16%
Total	0 (100%)	2 (100%)	30 (100%)	7 (100%)	6 (100%)	82 (100%)	18 (100%)	342 (100%)	100%

5.5 Discussion

5.5.1 Heterologous expression in yeast

The *ers1*Δ yeast strain is reported to have a hygromycin sensitive phenotype (Gao et al., 2005). However, despite several attempts no hygromycin hypersensitivity was observed in this study. The specific amino acid requirements of the strain (i.e. inability to grow on media without uracil, tryptophan, leucine and lysine) were observed as expected suggesting cross-contamination of strains is an unlikely explanation.

It is possible that the phenotype described in Gao et al. (2005) is affected by subtle differences in growth conditions between laboratories. As far as possible, identical conditions were used and care was taken to test a wide range of hygromycin B concentrations. Nevertheless, hygromycin sensitivity is affected by a variety of cellular processes (see below) and the exact cause for increased sensitivity in *ers1*Δ is unknown (Gao et al., 2005) making it difficult to ascertain the precise reason why the phenotype was not reproducible under our conditions.

However, despite the lack of hygromycin hypersensitivity in *ers1*Δ, a slight increase in hygromycin tolerance was seen when the human PQL protein CTNS was expressed in yeast. This is consistent with the results described in Gao et al. (2005). Differences in hygromycin tolerance can have a number of causes. Hygromycin is a cationic aminoglycosidic antibiotic that acts by inhibiting protein synthesis (Gonzalez et al., 1978). Mutations leading to altered hygromycin tolerance relate to a number of physiological processes including ion transport across the plasma membrane and tonoplast (see below), glycosylation (Dean, 1995) and vesicle trafficking (Conboy and Cyert, 2000).

In this case, given that CTNS is an endomembrane localised cystine/proton co-symporter (Kalatzis et al., 2001), the most likely scenario is that hygromycin tolerance results from cytosol acidification. As plasma membrane potential is known to directly influence hygromycin B (a cation) uptake (Perlin et al., 1988), cytosol acidification may result in depolarisation, reduced uptake and increased tolerance. A similar explanation

was used to explain the hygromycin sensitivity of *nhx1Δ*, a strain lacking an endomembrane localised cation-proton antiporter (Brett et al., 2005).

Assuming that *AtPQL* proteins are functional when heterologously expressed in yeast, it is surprising that none conferred hygromycin tolerance like CTNS. Either the *AtPQL* proteins are not functionally homologous to CTNS, less effective in conferring hygromycin resistance or not fully functional without plant specific post-translational modifications.

5.5.2 Isolation and physiological characterisation of mutant lines

Several different mutant lines were characterised in this study. Over-expressers were isolated for each family member and knockouts for each except *AtPQL2* (no T-DNA insertion mutants are available within the coding region of *AtPQL2*). The fact that homozygous lines could be found indicate that none of the mutations lead to gamete or embryo lethality. This is perhaps surprising given that some family members (*AtPQL1*, *AtPQL2* and *AtPQL5*) are preferentially expressed in pollen and others (*AtPQL4* and *AtPQL6*) are predicted to play a role in Man/Glc-P-Dol dependent glycosylation which appears to be essential for genetic transmission through pollen (see below). In addition, no phenotype at the whole plant level could be found under a variety of different conditions for any of the mutant lines.

However, a small reduction in the genetic transmission of both *pql4-1* and *pql4-2* knockout alleles was observed. Fewer *pql4* homozygotes were recovered from the self-fertilisation of a heterozygous parent plant than expected. However, more replicates are required to ascertain whether or not this affect is statistically significant and reciprocal crosses with wild type needed to determine whether the defect affects male-transmission, female transmission of both. Nevertheless, the reduced transmission of the knockout allele is consistent with the proposed function of *AtPQL4* in Man/Glc-P-Dol dependent glycosylation given that multiple mutants affected in this process (*sst3a/sst3b*, *gcs1*, *knf1*, *rsw3*, *pnt1*, *seth1* and *seth2*) are reported to have reduced gamete transmission (see chapter 1; table 1-1).

An obvious explanation for the lack of strong phenotypes at the whole plant level and in terms of genetic transmission is functional redundancy, particularly between *AtPQL4* and *AtPQL6* which are very similar to each other. It may be that mutation of multiple genes at the same time is required to see a phenotype at the whole plant or gamete level. However, *pql4-2/pql6-1* double mutants have recently been isolated and appear to be viable. The lack of phenotype may also be due to the *pql6-1* mutation being not completely null (although transcript level is greatly reduced).

The fact that no T-DNA insertion mutants are available within the coding region of *AtPQL2* is interesting. The gene does not appear to lie within a region with a low frequency of T-DNA insertion (multiple insertions are present within the coding region of adjacent genes). Although this could have occurred by chance, it could also be that the gene is essential for both male and female transmission making mutant recovery impossible.

5.5.3 Transcriptional profiling of mutant lines

Despite no obvious phenotype at the whole plant level, transcriptional profiling of *AtPQL4* and *AtPQL6* mutants revealed clear phenotypes at the molecular level. I will now discuss three themes identified from analysis of the microarray data. The first was differential regulation between mutants and wild type of genes and functional groups related to pathogen defence and cell-surface modification. The second was differential regulation of genes related to lipid metabolism and ER stress and the third was differential regulation of genes encoding secretory pathway associated and N-glycoproteins. All three can be explained by possible glycosylation defects.

5.5.3.1 Defence and cell-surface related genes

As highlighted in tables 5-5 and 5-6, multiple groups of genes encoding defence and cell-surface associated proteins showed differential regulation in each of the mutants profiled. This included genes encoding known pathogenesis related (PR) proteins, LTPs and cell-wall associated proteins.

PR genes: A number of *PR* genes were differentially regulated in the mutants compared to wild type. They include *PR-1*, which was up-regulated in both *pql6-1* and *35S::AtPQL6-1*; *PR-2*, which was up-regulated in *pql6-1*; *PDF1.2a* and *PDF1.2b* which were up-regulated in *pql6-1*; *THI2.2* and a putative thionin which were down-regulated in *pql6-1* and *pql4-2*; and several *LTPs* (see below).

PR proteins are a heterogeneous grouping of proteins that accumulate in response to pathogen attack. Various functions have been proposed for *PR* proteins ranging from direct anti-pathogen activity to normal developmental processes that occur independently of biotic stress (Van Loon and Van Strien, 1999). In general, *PR* proteins are highly adapted to resist the low pH and proteolytic conditions present in the vacuolar lumen and extracellular space where they are localised.

LTPs: The differential regulation of genes encoding *LTPs* and *LTP*-like proteins was one of the most interesting transcriptional responses unveiled by microarray analysis. iGA analysis showed that the lipid transport group (GO 0006869) had significantly increased expression in all three knockout lines and significantly decreased expression in *35S::AtPQL6-1*. This group primarily consists of *LTP* and *LTP*-like proteins including 2S seed storage proteins.

LTPs are small proteins (7-10 KDa) that are trafficked through the endomembrane system and secreted in the extracellular space where they accumulate in association with the cell wall (Salcedo et al., 2007). An initially proposed role in lipid shuffling between endomembrane compartments, based on the lipid-binding capacity of *LTPs*, seems unlikely given their extracellular localisation (Kader, 1996). However, several studies have implicated *LTPs* in plant defence (Salcedo et al., 2007). *LTP* gene expression is known to be induced by pathogens, and as a result *LTPs* have been classified as a *PR* family (*PR*-14; Van Loon and Van Strien, 1999). Purified *LTPs* have been reported to have anti-pathogenic properties *in vitro* (Molina et al., 1993) and transgenic *A. thaliana* or tobacco plants over-expressing a barley *LTP* show increased tolerance to bacterial pathogens (Molina and Garcia-Olmedo, 1997).

LTPs have also been proposed to act as carriers in the transport of wax precursors from the interior of epidermal cells to the developing cuticle (Yeats and Rose, 2008). In

accordance with this hypothesis, increased LTP expression has been correlated with increased wax accumulation (Hollenbach et al., 1997; Cameron et al., 2006).

Cell-wall biosynthesis and expansion: A number of genes associated with cell wall modification were highlighted by iGA and RP. Particularly prominent was *XTR7*, which encodes a putative xyloglucan endotransglycosylase (XET) and was the only gene other than *AtPQL6* itself that showed opposite regulation in *pql6-1* and *35S::AtPQL6-1*. *XTR7* was up-regulated in *pql6-1* and down-regulated in *35S::AtPQL6-1*. XETs are cell wall modifying enzymes that cleave and rejoin xyloglucan chains and are associated with cell wall expansion (Campbell and Braam, 1999). XETs have also been implicated in the response to pathogens. One *A. thaliana* XET, *XTH33* is up-regulated by pathogen attack and *xth33* mutants are preferentially targeted by aphid herbivores (Divol et al., 2007).

Another interesting finding of iGA was the up-regulation of the *COBRA* family in *pql6-1*. One family member, *COBL2*, is also on the RP list of up-regulated genes in *pql6-1*. *COBRA* itself encodes a GPI anchored and N-glycosylated cell wall associated protein required for correct cellulose microfibril organisation and directionally dependent control of cell wall expansion (Roudier et al., 2005). *COBL2* is yet to be functionally characterised but is interesting in that it is one of only a few family members reported to show strong stress induction at the transcript level (Brady et al., 2007).

Links between glycosylation and defence: The fact that so many defence-related genes were differentially regulated in mutants with predicted defects in Man/Glc-P-Dol utilisation, suggests a link between glycosylation and defence. There have been several previous reports connecting the two processes. Firstly, Wang et al. (2005) showed that expansion of the protein secretory pathway is induced and required for systemic acquired resistance (SAR). They identified several defence and secretory pathway related genes as primary targets of the NPR1 transcription factor (a key regulator of SAR). Up-regulated genes included *BiP2*, *DAD1*, *PMM* and two genes each encoding calnexin and calreticulin isoforms all of which are involved in N-glycosylation (see chapter one for description of functions). Wang et al. (2005) subsequently showed that *bip2* and *dad1* mutants failed to secrete PR1 in response to infection and concluded that an increase in secretory pathway capacity is an important component of SAR.

Furthermore, during pathogen attack *BiP* transcripts are induced prior to PR1, presumably in anticipation of a requirement to increase secretory pathway capacity (Jelitto-Van Dooren et al., 1999). In addition, chemical inhibitors of N-glycosylation have been shown to induce SAR in cucumber (Sticher and Metraux, 2000) and the hypersensitive response in tobacco cell cultures (Iwata and Koizumi, 2005a). One explanation is that SAR and the hypersensitive response are secondary effects caused by tunicamycin-induced cell death. However, the relationship between cell death and SAR is unlikely to be causal and other stresses leading to cell death do not necessarily induce SAR (Greenberg, 1997).

5.5.3.2 Lipid-metabolism and ER stress-related genes

Given the potential role of *AtPQL4* and *AtPQL6* in N-glycan biosynthesis in the ER, it was hypothesised that mutants might show an ER stress response. The transcriptional profiles of *AtPQL4* and *AtPQL6* gave strong indications of differential expression of genes associated with two components of the ER stress response – the unfolded protein response (UPR) and the inositol response (Cox et al., 1997). The UPR was most apparent in *AtPQL6* mutants where RP analysis showed differential regulation of multiple UPR-associated genes. In both *AtPQL4* and *AtPQL6* mutants, there was also strong evidence of differential regulation of genes associated with lipid biosynthesis and metabolism which may form part of the inositol response.

Unfolded Protein Response (UPR): One of the most interesting findings uncovered by microarray analysis and confirmed by qPCR was up-regulation of *BiP3* in *pql6-1* and *35S::AtPQL6-1*. *BiP3* is one of three *BiP* genes in *A. thaliana* encoding ER localised luminal binding proteins (Noh et al., 2003). BiPs preferentially bind under-glycosylated and incorrectly folded proteins in the ER lumen. In yeast, BiPs have been described as master regulators of the UPR, which activate ER stress signal transducers in response to an accumulation of unfolded proteins (Zhang and Kaufman, 2004).

In plants (as well as other eukaryotes), the up-regulation of *BiP* genes is a typical response to disruptions in N-glycan biosynthesis. In *A. thaliana*, tunicamycin induces the expression of all three *BiP* genes (Iwata and Koizumi, 2005b). Similarly, mutants involved in N-glycan trimming and the calnexin-calreticulin cycle have been reported to show increased *BiP* transcript level (Taylor et al., 2000). Furthermore, over-expression of *BiP* genes in tobacco confers tolerance to tunicamycin (Leborgne-Castel et al., 1999; Alvim et al., 2001).

Another UPR-related gene, *AtUTr1* was up-regulated in *35S::AtPQL6-1*. *AtUTr1* encodes an ER-localised nucleotide sugar transporter (see chapter 1, section 1.3.2.4) that provides UDP-Glc for the re-glucosylation of incorrectly folded N-glycoproteins (see chapter 1, section 1.3.2.4 ; Reyes et al., 2006). By doing so it encourages interaction of newly synthesised proteins with calnexin and calreticulin and the establishment of correct folding before ER export. *AtUTr1* was also identified as up-regulated in response to tunicamycin in three previous microarray studies (Martinez and Chrispeels, 2003; Noh et al., 2003; Kamauchi et al., 2005).

In addition, four *PPI* genes, encoding known or putative peptidyl prolyl cis trans isomerases (PPIases), were up-regulated in *35S::AtPQL6-1*. PPIases play an important role in protein folding by catalyzing the isomerisation of peptide bonds immediately before and after proline residues (Wang and Heitman, 2005). Whereas the majority of peptide bonds occur in the energetically favourable trans form, peptide bonds involving proline may also occur in the cis form and isomerisation to this form is important for protein folding. In animals, PPIases are well known components of the UPR and are induced at the transcriptional level by ER stress (Kultz, 2005). One of the *PPI* genes up-regulated here, *ROF1*, is also known to be induced by heat shock (Aviezer-Hagai et al., 2007).

Inositol Response: Several genes encoding enzymes involved in lipid metabolism were found to be differentially regulated in *AtPQL* mutants. It is possible that these changes constitute an inositol response complimentary to the UPR. The inositol response involves a wide variety of different lipid metabolic enzymes that function to increase ER capacity under stress (Stroobants et al., 1999). In plants, the inositol response is poorly understood. Tunicamycin has been reported to increase phosphatidylinositol 4-kinase

activity in *Glycine max* cell cultures (Shank et al., 2001) but previous transcriptional profiling studies identified few genes involved in lipid metabolism with altered expression under ER stress (Martinez and Chrispeels, 2003; Noh et al., 2003; Kamauchi et al., 2005).

Here, several lipid-metabolism related genes were identified as differentially regulated. The mutants with the most noticeable changes in lipid metabolism-related gene regulation were *pql4-2* and *35S::AtPQL6-1*. In both mutants, iGA recognised a group of two genes with C22 sterol desaturase activity (CYP710A1 and CYP710A2) as down-regulated relative to wild type. Genes with predicted function in acyl lipid metabolism were also down-regulated in both mutants.

Furthermore, both mutants showed up-regulation of other groups involved in lipid metabolism. This includes groups annotated as seed oily body biogenesis in *pql4-2* and a group involved in isopentenyl diphosphate biosynthesis in *35S::AtPQL6-1*. Also of interest is up-regulation of a group encoding proteins with phosphoethanolamine-N-methyltransferase activity in *pql4-2*. These enzymes catalyze a key step in the production of choline, which is in turn the precursor of phosphatidylcholine, which accounts for 40-60% of the total lipid content of non-plastid plant membranes (Mou et al., 2002).

In addition, a *DPS-like* gene was down-regulated in *pql4-2*. This gene is homologous to *DPS*, the sole *A. thaliana* dehydrolipichol pyrophosphate (dedol-P-P) synthase characterised to date (Cunillera et al., 2000). The differential regulation of this gene is of particular interest given that dedol-P-P synthase catalyzes the production of dedol-P-P, the precursor to Dol-P-P, the lipid acceptor of sugar residues during N-glycan biosynthesis (see chapter 1, section 1.3.2.1).

5.5.3.3 Secretory pathway-associated and N-glycoprotein encoding genes

The over-representation of genes encoding secretory pathway and/or N-glycosylated proteins amongst up-regulated genes in *pql4* and *pql6* knockouts may reflect impaired glycosylation. If in the mutants less N-glycoproteins are being correctly processed then the genes encoding them may be differentially regulated in order to compensate. For example, if only 50% of a particular N-glycoprotein is being correctly

folded and processed then it may be that twice as much transcript is required to achieve the same level of function. The over-representation of differentially regulated genes encoding proteins targeted to the secretory pathway probably reflects that many of these proteins are N-glycosylated. In addition, even proteins which are not N-glycosylated could be affected by the extra workload imposed upon the ER by glycosylation defects.

The fact that the opposite is seen amongst up-regulated genes in *35S::AtPQL6-1* is interesting (under-representation of secretory pathway and N-glycoproteins). One explanation is that N-glycosylation is more efficient in the over-expresser and that therefore less transcript is required (see below).

5.5.3.4 Conclusions

The microarray experiments provided evidence for induction of an unfolded protein response (UPR), a possible induction of an inositol response and differential regulation of known defence-related, glycoprotein-encoding and secretory pathway associated genes in the mutants (especially *AtPQL6* mutants). Figure 5-23 provides a possible explanation for these responses. I predict that reduced expression of *AtPQL4* and *AtPQL6* directly effects Man/Glc-P-Dol utilisation. As is the case with the *lec35* mutants of CHO cells (Lehrman and Zeng, 1989), *AtPQL4* and *AtPQL6* mutants may fail to fully extend the lipid linked oligosaccharide precursor during N-glycan biosynthesis and, as a result, transfer Man₅GlcNAc₂ rather than Glc₃Man₉GlcNAc₂ to nascent N-glycoproteins. Man₅GlcNAc₂ would presumably then be mono-glucosylated by the enzyme UDP-glucose:glycoprotein glucosyltransferase producing Glc₁Man₅GlcNAc₂ whereas wild type cells produce Glc₁Man₉GlcNAc₂ following Glc trimming (see chapter 1 section 1.3.2.4). This difference is likely to impair interaction of newly synthesized glycoproteins with the ER folding chaperones, calnexin and calreticulin. *In vitro* binding assays have shown that calnexin has a much greater affinity for Glc₁Man₉GlcNAc₂ than Glc₁Man₅GlcNAc₂ (Ware et al., 1995).

An impaired interaction with calnexin and calreticulin is likely to increase the amount of misfolded proteins in the ER lumen and trigger the ER stress response. A build up of unfolded proteins may be sensed by BiP3, which is known to be a key regulator of

the ER stress response. Not only do BiP proteins have intrinsic chaperone activity themselves, they also play a role in signal transduction to the nucleus (Zhang and Kaufman, 2004). In this case, the increased expression of *BiP3* may be a signal for up-regulation of other components of the UPR and inositol response but its lack of induction in the *pql4* knockouts suggests it is not the only signal.

Another key finding was the differential regulation of several genes encoding N-glycoproteins and defence and cell-surface related proteins. It is predicted that glycosylation defects in *AtPQL* mutants reduce the capacity of the secretory pathway to traffic proteins to the extracellular space. It is possible that feedback mechanisms exist to maintain the amount of secreted proteins (many of which are defence related) by increasing transcription of the genes responsible.

There is some similarity between the results presented here and those of Sticher and Metraux (2000) who reported a triggering of systemic acquired resistance in cucumber following treatment with tunicamycin and amphomycin (both are inhibitors of N-linked glycosylation – see chapter 1, figure 1-2, tunicamycin acts at step 1 and amphomycin at step 9). In both cases defects in N-glycosylation and a reduction in the amount of proteins reaching the cell surface may have led to the up-regulation of defence-related genes.

The fact that N-glycoprotein encoding genes were under-represented amongst up-regulated genes in *35S::AtPQL6-1*, yet over-represented amongst up-regulated genes in the knockouts is a classical example of an opposite phenotype in knockouts and over-expressers and is in accordance with the compensatory mechanism proposed above. However, the observation that both *35S::AtPQL6-1* and *pql6-1* mount an ER stress response (i.e. the same phenotype) may at first glance seem confusing. In this respect, *35S::AtPQL6-1* could act as a loss of function mutant. It may be that both *AtPQL4* and *AtPQL6* are required for Man/Glc-P-Dol utilisation but when *AtPQL6* is over-expressed it out-competes *AtPQL4* in protein complexes and/or lipid sub-domains. As a result *AtPQL6* over-expression may have a dominant negative effect on *AtPQL4* function. A similar explanation was used to explain the effect of *GPT* over-expression in CHO cells, which leads to a *lec35* like phenotype (see chapter 1, section 1.3.7.4).

Therefore, in *35S::AtPQL6-1* mutants the ER stress response and the differential regulation of N-glycoproteins may be triggered by distinct mechanisms. Whereas the ER stress response is likely due to an increase in misfolded proteins, the under-representation of N-glycoprotein encoding genes may be triggered by other effects of over-expression up-stream of protein misfolding. For example, *AtPQL6* itself may interact with ER-resident signalling components to alter the transcription of N-glycoprotein encoding genes.

Figure 5-23. Proposed explanation for the molecular phenotype of *AtPQL4* and *AtPQL6* mutants. Black text indicates experimental evidence; blue indicates predictions based on microarray results and literature.

A. *pql4-1* and *pql4-2* have no expression of *AtPQL4*, and *pql6-1* has decreased expression of *AtPQL6*. *35S::AtPQL6-1* may act as a loss-of-function dominant negative mutant with *AtPQL6* replacing *AtPQL4* in lipid sub-domains and/or protein complexes.

B. *AtPQL4* and *AtPQL6* are homologous to LEC35/MPDU1 which is essential for the utilisation of Man/Glc-P-Dol. As a result of their reduced expression, Man/Glc-P-Dol utilisation may be decreased and Man₅GlcNAc₂ attached to newly synthesised proteins rather than Glc₃Man₉GlcNAc₂.

C. Interaction of newly synthesised N-glycoproteins with the ER folding chaperones, calnexin and calreticulin, may decrease. Calnexin and calreticulin are known to bind poorly to Glc₁Man₅GlcNAc₂.

D. As a result of decreased interaction with calnexin and calreticulin, the amount of incorrectly folded proteins in the ER lumen is likely to increase.

E. Increased protein misfolding in the ER is the probable trigger of ER stress and the up-regulation of *BiP3* which acts as a sensor for unfolded proteins. *BiP3* also has intrinsic chaperone activity and may reduce the amount of protein misfolding.

F. *AtPQL* mutants showed up-regulation of components of the UPR (PPIases/UTr1) and differential regulation of genes involved in lipid metabolism. The trigger for this may be *BiP3* or other signalling components associated with ER stress. Again, the up-regulation of these genes may reduce protein misfolding.

G. As a result of glycosylation defects, the number of proteins reaching the cell surface may decrease. This may include lipid transfer proteins, defence related proteins, and cell wall associated proteins, many of which are heavily glycosylated.

H. It was found that several genes encoding proteins targeted to the secretory pathway (such as LTPs, defence-related and N-glycosylated proteins) were differentially regulated in *AtPQL4* and *AtPQL6* mutants. The trigger for this may be ER stress (H), a reduction in the number of proteins reaching the cell surface (G), or a direct effect of *AtPQL* misexpression (A).

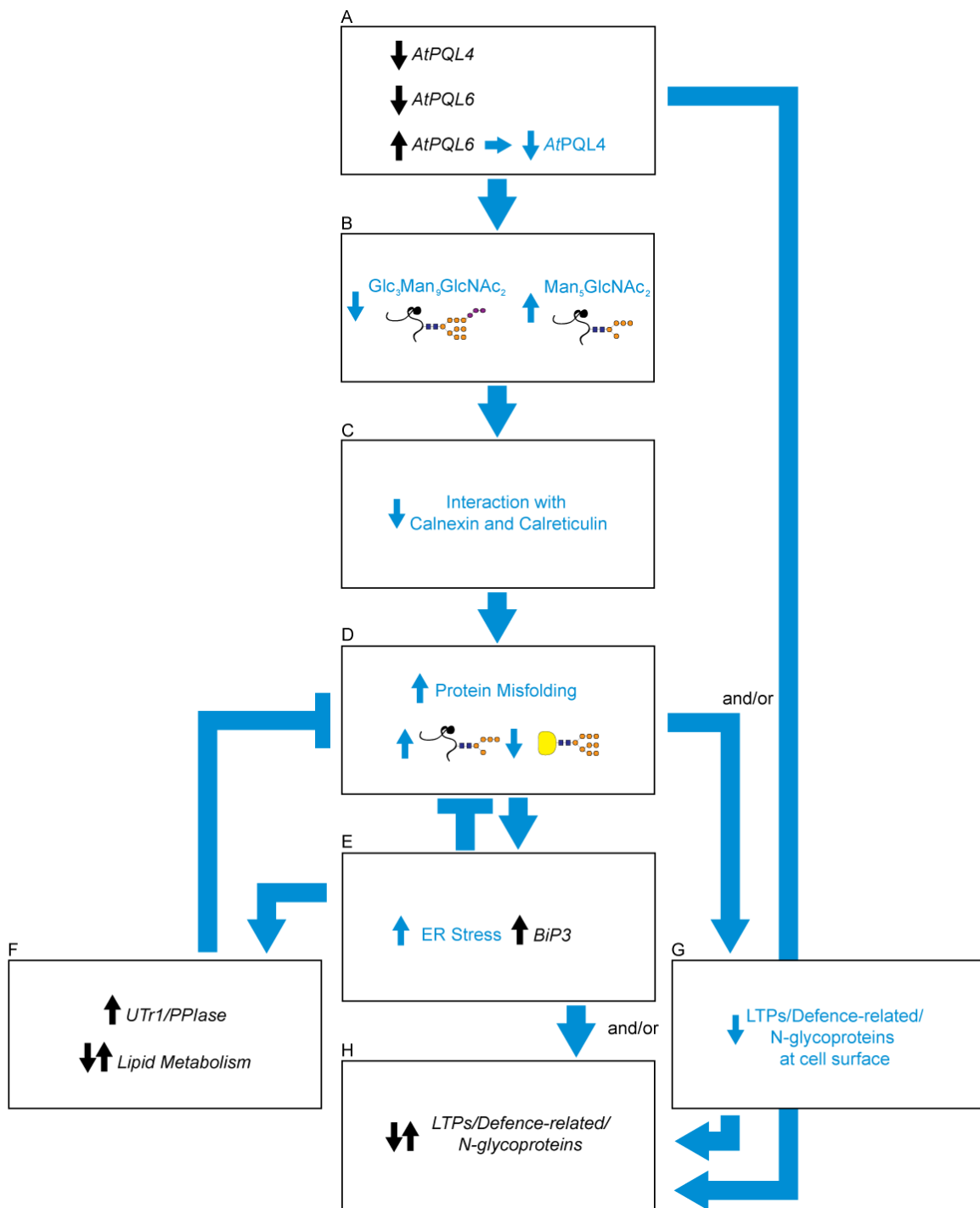


Figure 5-23. Proposed explanation for molecular phenotype of *AtPQL4* and *AtPQL6* mutants. See opposite for full legend.

Chapter 6: General Discussion

The central aim of this thesis was to provide a thorough characterisation of the *AtPQL* family. This has been achieved at the level of sequence analysis and through a variety of experimental approaches. Table 6-1 provides a summary of the main points made in this thesis. Discussion of individual experiments is made in each chapter. The purpose of this final discussion is to identify the important themes, drawing together information from different chapters to provide an overall picture of the *AtPQL* family. In particular I will discuss the likelihood of functional homology between the *AtPQL* proteins and their closest homologues, and discuss common threads running through the family. Finally, I will pose open questions which remain to be answered and make suggestions for further work.

6.1 Assessment of *AtPQL* function in plants

6.1.1 Are any *AtPQL* proteins functionally homologous to STM1?

STM1 is a PQL protein and putative GPCR from *S. pombe* (Chung et al., 2001). Although GPCR activity has not yet been categorically confirmed, there is clear evidence of physical and genetic interaction with the *S. pombe* G α , GPA2. The closest homologues of STM1 in *A. thaliana* in terms of sequence similarity are *AtPQL1*, *AtPQL2* and *AtPQL3*. The four proteins cluster together on a phylogenetic tree of the PQL family and share a common organisation of trans-membrane and PQ-loop domains.

It is the conclusion of this thesis that functional homology between *AtPQL1-3* and STM1 is very unlikely. If *AtPQL1-3* were functionally homologous to STM1 and acted as GPCRs it would be expected that they would interact with GPA1, the sole G α subunit in *A. thaliana*. However, key residues conserved between STM1 and other known or putative GPCRs that have been identified as critical for STM1-G α binding are not conserved in *AtPQL1-3* (see chapter 3, figure 3-7b)

Furthermore, according to the classical model of G-protein signalling, GPCRs localise to the plasma membrane to allow ligand binding and G α interaction. For at least one of *AtPQL1-3*, this does not appear to be the case. There is good evidence that

AtPQL1 is localised to the tonoplast, inconsistent with function as a GPCR.

Finally, it would be expected that *AtPQL1-3* mutants would have overlapping (or opposite) phenotypes with *GPA1* mutants. However, no phenotype was found for any of the *AtPQL1-3* mutants isolated in this study.

6.1.2 Are any *AtPQL* proteins functionally homologous to CTNS?

CTNS is a human lysosomal cystine transporter that is essential to prevent cystine hyper-accumulation as a result of lysosomal protein hydrolysis (Kalatzis et al., 2004). The closest *A. thaliana* homologue of CTNS is *AtPQL5*. A functional homologue of CTNS would be expected to not only prevent hyper-accumulation of cystine in lysosomes but also contribute to cysteine (cys) homeostasis. Consistent with a similar role in preventing cystine hyper-accumulation, *AtPQL5* is localised to the vacuole, the main site of protein hydrolysis in plants (Callis, 1995). However, the lack of an obvious phenotype in *pql5-1* suggests that *AtPQL5* is not essential to prevent cystine hyper-accumulation under normal conditions. In addition, the lack of a phenotype under several different stresses (including some where sulfur was limited) argue against *AtPQL5* playing a CTNS-like function during cys homeostasis. Although this does not necessarily prove that *AtPQL5* is not functionally homologous to CTNS, it does show that unlike CTNS it is dispensable for normal growth and development under a variety of adverse conditions.

6.1.3 Are any *AtPQL* proteins functionally homologous to MPDU1?

MPDU1/LEC35, is a mammalian PQL protein known to be essential for the utilisation of Man/Glc-P-Dol as a sugar donor during N-glycan and GPI anchor biosynthesis in the ER (Anand et al., 2001). The closest *A. thaliana* homologues of MPDU1/LEC35 are *AtPQL4* and *AtPQL6*. The proteins have a similar organisation of PQ-loops and transmembrane domains and two key residues known to be essential for MPDU1 function (a glycine at position 73 and a leucine at position 74) are conserved in both *AtPQL4* and *AtPQL6* (see chapter 3, table 3-3)

Both *AtPQL4* and *AtPQL6* appear to be localised to the ER and are expressed in a variety of different tissues. This is consistent with a homologous role to MPDU1/LEC35

as the ER is the site of N-glycan and GPI-anchor biosynthesis and other genes involved in these processes are reported to be expressed widely (see chapter 4, section 4.3.1.2).

Further evidence of a role in glycosylation is seen in the transcriptional profiles of *AtPQL4* and *AtPQL6* mutants. The differential regulation of several genes associated with the response to ER stress and others encoding secretory pathway associated proteins and N-glycoproteins is highly indicative of a defect in N-glycosylation.

However, despite the differences in gene expression, no developmental phenotype (except a small reduction in fertility in *pql4-1* and *pql4-2*) was seen for any of the *AtPQL4* or *AtPQL6* mutants. This is surprising given the severe phenotype of other mutants involved in N-glycosylation and GPI anchor biosynthesis (see chapter 1, table 1-1). An obvious explanation is functional redundancy between *AtPQL4* and *AtPQL6*. However, *pql4-2/pql6-1* double mutants have recently been isolated and appear to be viable. Alternatively, as *pql6-1* does not completely abolish *AtPQL6* transcript there may be residual activity. The changes in gene expression unveiled by microarray analysis may also compensate for the mutation. For example, the altered expression of several N-glycoproteins might make up for some of them being incorrectly glycosylated.

6.1.4 Speculative role in proton translocation

The PQL proteins characterised to date have apparently diverse functions. This may suggest that their defining characteristic, the repeated PQ-loop, is solely a structural feature connected to protein folding with few implications for function.

However, there is some evidence of conservation of function even between family members with relatively low sequence homology. One possible common feature in all the PQL proteins described so far is the coupling of their primary activity to the movement of protons.

The evidence is clearest for CTNS, which is known to couple the transport of cystine to the movement of protons (Kalatzis et al., 2001). In addition, yeast strains lacking ERS1, have been reported to have a hygromycin sensitive phenotype which can be complemented by CTNS despite the fact that sequence similarity between the mutants is relatively low and the proteins do not cluster together on phylogenetic trees of the PQL family (Gao et al., 2005). A possible cause of the hygromycin hypersensitive phenotype

of this strain is defects in pH homeostasis (see chapter 5, section 5.7.1). However, it should be noted that the reported phenotype of this line was not reproducible in this study. In addition, it is proposed that MPDU1/LEC35 may be involved in the flippase-mediated transport of Man/Glc-P-Dol from the cytosol to the ER lumen (Anand et al., 2001). Given that re-orientation of Man/Glc-P-Dol across lipid membranes is thermodynamically highly unfavourable yet ATP-independent (Rush and Waechter, 1995, 1998) one possible way to provide energy for this process would be to couple the flippase activity to the movement of protons. Although the role MPDU1/LEC35 plays in this process is unknown, and there is some evidence that it does not itself have flippase activity itself (Anand et al., 2001), it may be involved in providing energy to the flippase.

Further support for this hypothesis comes from homologues of the PQL family. Firstly, both characterised members of the related TDT family (see chapter 3, figure 3-9), are proton-coupled transporters (Grobler et al., 1995; Turner et al., 1997) and secondly the microbial rhodopsins, previously shown to be distantly related to the PQL family, also translocate protons (Zhai et al., 2001).

6.1.5 Up-regulation by wounding and viral infection

One of the surprising findings of this study was the up-regulation of four *AtPQL* genes (*AtPQL3-6*) by viral infection and wounding, and that this induction was enhanced in the *coil* mutant background. This data, along with the transcriptional profiles of *AtPQL4* and *AtPQL6* mutants, is suggestive of a role of the *AtPQL* family in defence. This is perhaps not surprising for *AtPQL4* and *AtPQL6* given the previous reports linking glycosylation in the secretory pathway with defence signalling (Wang et al., 2001; see chapter 4, section 4.3.2.1). However, for the others it is not as easy to infer what role they might play in defence.

One possibility is that their induction is linked to proton transport (see above). All four may be induced to contribute to cytosolic acidification. It is known that cytoplasmic pH is rapidly reduced upon treatment with fungal elicitors (Kuchitsu et al., 1997; He et al., 1998; Lebrun-Garcia et al., 2002), and it has been proposed that cytosolic acidification is an important part of the wound response, inducing the expression of several wound responsive-genes (Hara et al., 2000).

Table 6-1. Summary of proposed localisation, function and regulation of *AtPQL* proteins. In each column evidence for and against the given statement is given in *italics*. The table is colour coded to reflect the strength of the evidence for and against.

						<div>very unlikely (evidence against)</div> <div>likely (strong prediction/some experimental evidence)</div> <div>very likely (strong evidence in favour)</div>
Protein	Sequence Features	Closest Homologue	Tissue Localisation	Subcellular Localisation	Function	Regulation
<i>AtPQL1</i>	7 TMS, 2 PQ-Loops	STM1	Pollen (early stage) <i>qPCR, microarray</i>	Vacuole - <i>GFP</i>	<i>GPCR - FOR: similarity with STM1. AGAINST: incorrect membrane localisation, no phenotype, does not share crucial residues</i>	
<i>AtPQL2</i>	7 TMS, 2 PQ-Loops	STM1	Pollen - <i>qPCR, microarray</i>	Vacuole - <i>similarity to AtPQL1</i>	<i>GPCR - FOR: similarity with STM1. AGAINST: incorrect membrane localisation, no phenotype, does not share crucial residues</i>	
<i>AtPQL3</i>	7 TMS, 2 PQ-Loops	STM1	Roots, Floral Buds <i>qPCR</i>	Vacuole - <i>similarity to AtPQL1</i>	<i>GPCR - FOR: similarity with STM1. AGAINST: incorrect membrane localisation, no phenotype, does not share crucial residues</i>	Up-regulated by wounding and viral Infection only in <i>coi1</i> mutant background - <i>qPCR</i>
<i>AtPQL4</i>	7 TMS, 2 PQ-Loops	MPDU1	Ubiquitous - <i>qPCR</i>	ER - <i>GFP</i>	Man/Glc-P-Dol utilization - <i>FOR: close similarity to MPDU1, ubiquitously expressed, ER localized, differential expression of several N-glycoprotein encoding, secretory pathway associated and ER stress-related genes in mutant lines AGAINST: no developmental phenotype</i>	Diurnally Regulated (peak expression in evening) - <i>qPCR</i>
						Up-regulated by wounding and viral Infection. Induction enhanced in <i>coi1</i> mutant background - <i>qPCR</i>
<i>AtPQL5</i>	7 TMS, 2 PQ-Loops	CTNS	Pollen (mature), also relatively strong throughout plant <i>qPCR, Microarray</i>	Vacuole - <i>GFP</i>	Cystine export from vacuoles - <i>FOR: similarity to CTNS, vacuolar localisation AGAINST: No increased hygromycin tolerance in yeast, mutants show no developmental phenotype even on low sulphur medium</i>	Diurnally regulated (peak expression during day) - <i>qPCR, microarray</i>
						Up-regulated by wounding and viral Infection. Induction enhanced in <i>coi1</i> mutant background - <i>qPCR</i>
<i>AtPQL6</i>	7 TMS, 2 PQ-Loops, Evening element and CCA1 promoter motifs	MPDU1	Ubiquitous - <i>qPCR, microarray</i>	ER - <i>GFP</i>	Man/Glc-P-Dol utilization - <i>FOR: close similarity to MPDU1, ubiquitously expressed, ER localized, differential expression of several N-glycoprotein encoding, secretory pathway associated and ER stress-related genes in mutant lines AGAINST: no developmental phenotype</i>	Diurnally regulated (peak expression in evening)- <i>qPCR, microarray, promoter motifs</i>
						Up-regulated by wounding and viral Infection. Induction enhanced in <i>coi1</i> mutant background - <i>qPCR</i>

6.1.6 Prospects for further study

Of all the relationships between *AtPQL* proteins and their non-plant homologues, the similarity between *AtPQL4*, *AtPQL6* and MPDU1/LEC35 is perhaps the most informative. Unlike the other characterised PQL proteins, MPDU1/LEC35 plays a role in a key metabolic process known to be well conserved throughout the eukaryote kingdom (see introduction). Although its precise function is not fully understood, it is known to be essential. Whereas *A. thaliana* could conceivably have evolved without a functional homologue to CTNS and STM1, it seems less likely that it would not have a functional homologue of MPDU1.

Because of this and the experimental data presented here showing sub-cellular localisation, tissue specific distribution and mutant transcriptional profiles consistent with a role in Man/Glc-P-Dol dependent glycosylation, *AtPQL4* and *AtPQL6* are considered the most promising candidates for further study. Further study of these proteins should test for functional homology with MPDU1/LEC35 and look in detail at the role of these proteins in glycosylation. More broadly, the transcriptional profile of these mutants adds to a growing body of evidence suggesting a link between defence and glycosylation which may warrant further investigation.

Complementation of *lec35*: To test for functional homology between *AtPQL4/AtPQL6* and LEC35/MPDU1, the simplest approach may be to express the plant genes in the *lec35* CHO mutant cell line and test for complementation. The *lec35* mutant is more tolerant to a combined ConA/N-glycoprotein processing inhibitor treatment than wild type (see chapter 1, section 1.3.7.1) (Lehrman and Zeng, 1989). An alternative might be to express *LEC35* (or its human homologue *MPDU1*) in *pql4/pql6* mutant plants to check for complementation of the transcriptional phenotype (although this may be complicated by dose dependent effects).

To confirm homologous function *in planta*, direct analysis of lipid-linked oligosaccharides (LLOs; the precursor for N-glycosylation) should be undertaken. The *lec35* mutant accumulates Man₅GlcNAc₂-P-P-Dol instead of Glc₃Man₉GlcNAc₂-P-P-Dol (Lehrman and Zeng, 1989). As LLOs are present in very low quantities (approximately 1 nmol per gram in mammalian pancreatic tissue) direct analysis by HPLC/MS/TLC is

problematic (Gao and Lehrman, 2002). Traditionally, as was the case for the *lec35* mutants, LLOs have been studied by prior labelling with [³H]-Man. Although useful for cell cultures and isolated membranes, a drawback of this approach is that it is not readily applicable to whole organisms due to pool dilution, catabolism and inefficient distribution of the radiolabel (usually provided as GDP-[³H]-Man in the culture medium) (Gao and Lehrman, 2002).

An alternative approach is to use FACE (fluorophore assisted carbohydrate electrophoresis) to visualise LLOs separated according to size (Gao and Lehrman, 2002). Using this technique, intact LLOs are extracted and the dolichol molecule removed by mild acid hydrolysis. The sugars are then labelled with a fluorophore, separated by polyacrylamide gel electrophoresis and compared to oligosaccharide standards.

Alternatively rather than analysing the LLO precursors, it may be possible to analyze glycoproteins directly, either through HPLC/TLC/MS or protein blotting techniques using probes for specific sugar residues. However, such approaches are complicated by the fact that the *lec35* mutation affects mannosylation of LLOs only after the production of Man₅GlcNAc₂-P-P-Dol (Zeng and Lehrman, 1990). If *pql4* and *pql6* mutants have the same defect, glycoproteins would still retain the pentasaccharide core glycan (Man₃GlcNAc₂). In wild type cells maturation to complex glycans only occurs after ER and Golgi mannosidases and glucosidases trim glycans back to this pentasaccharide core (Strasser et al., 2007). Therefore, we predict that *pql4* and *pql6* retain the ability to produce both high mannose and complex glycans identical to wild type. However, a difference may be seen if the plants are first treated with the golgi mannosidase II inhibitor, swainsonine, which has been shown to prevent the production of complex glycans (Tulsiani et al., 1982; Mega, 2004). As has been suggested for *lec35* (Lehrman and Zeng, 1989), we predict that *pql4* and/or *pql6* mutants may effectively bypass golgi mannosidase II and ER-resident glucosidases and unlike wild type, continue to produce complex glycans in the presence of swainsonine.

Another option would be to take a genetic approach. One option would be to make double/triple mutants with *pql4* and/or *pql6* and glucosidase I mutants (*knf1/gcs1*, see chapter 1, section 1.3.2.4). *gcs1* and *knf1* both have embryo lethal phenotypes as a result of being unable to trim the two terminal Glc residues from Glc₃Man₉GlcNAc₂. If

AtPQL4 and/or *AtPQL6* plays a homologous role to LEC35/MPDU1 then these Glc residues would never be added, and knocking out *AtPQL4* and/or *AtPQL6* should rescue the embryo lethal phenotype of the glucosidase I mutants.

Similarly, double/triple mutants could be made with *pql4* and/or *pql6* and the golgi mannosidase II mutant, *hgl1-1*. *hgl1-1* is unable to convert hybrid type glycans to complex glycans (Strasser et al., 2006) and has in effect the same phenotype as swainsonine treated plants. I predict that knocking out *AtPQL4* and/or *AtPQL6* in the *hgl1-1* background should be able to complement this phenotype.

Role of LLO synthesis in ER stress and pathogen defence: The link between N-glycan biosynthesis and ER stress is well established and significant progress has been made in identifying important signalling components (see chapter 1, section 1.3.2.2). However, little is known about LLO flux during ER stress in plants. In animals, control of LLO synthesis under ER stress is a priority and cellular processes are modified to ensure Glc₃Man₉GlcNAc₂-P-P-Dol levels are maintained reflecting its role as a key folding determinant (Shang et al., 2007). It is known that Glc₃Man₉GlcNAc₂-P-P-Dol is the preferred substrate for transfer to newly synthesised glycoproteins, and glycoproteins with incomplete oligosaccharide chains (lacking Man or Glc residues) interact poorly with ER folding chaperones (Ware et al., 1995). In animals, a minimum level of Glc₃Man₉GlcNAc₂-P-P-Dol is achieved by decreasing protein synthesis to reduce consumption and increasing the rate of conversion from LLO intermediates (Shang et al., 2007).

The adaptation of FACE protocols could facilitate the investigation of LLO flux in plants which may have implications for the response to multiple stresses that can adversely affect protein folding (e.g. salt stress, heat shock, dehydration etc.) and increase pressure on the ER (e.g. pathogen attack – through the increased synthesis of secreted proteins). Several reports indicate that genes encoding proteins known or predicted to be involved in the production or transport of LLO intermediates are stress inducible including *PMM* (Wang et al., 2005), *AtUTr1* (Reyes et al., 2006) and *GPT* (Koizumi et al., 1999). It would be interesting to see how tightly controlled LLO flux is in plants, whether or not it is modified in response to stress, and whether it can be engineered in order to increase stress tolerance. For example, if LLO flux could be engineered to

increase Glc₃Man₉GlcNAc₂-P-P-Dol synthesis, perhaps by over-expressing ER glycosyltransferases, this may reduce protein misfolding and increase stress tolerance.

In addition, there is increasing evidence of a link between N-glycosylation and defence. It has been shown that NPR1, a key transcription factor involved in systemic acquired resistance, is responsible for the induction of several secretory pathway, and especially N-glycosylation, related genes (Wang et al., 2005). Not only were genes encoding chaperones and proteins with a direct effect on protein folding up-regulated, but also genes encoding proteins such as phosphomannomutase and *AtUTr1* which have direct effects on substrate availability for LLO synthesis. In addition BiP genes, which encode key regulators of the ER-stress response, are induced by pathogen attack before classical defence proteins such as PR-1 (Jelitto-Van Dooren et al., 1999). These results suggest that one of the early cellular responses to pathogen attack is expansion of the secretory pathway. This is most likely in anticipation of ER stress due to an increased requirement to synthesise defence related proteins, many of which are secreted.

Furthermore, inhibition of LLO synthesis by tunicamycin and amphomycin induces systemic acquired resistance (Sticher and Metraux, 2000). This, in combination with the evidence presented here, that probable alteration of LLO flux by mutagenesis induces defence-related genes, suggests co-ordinated regulation of N-glycosylation and defence responses.

Given these links, a clear understanding of the mechanisms by which secretory pathway capacity and LLO flux are controlled is desirable. The interaction between pathogen defence and N-glycosylation may have important implications for the design of pathogen-resistant transgenic plants. It is interesting to consider whether plants could be manipulated to increase ER capacity and optimise LLO flux, and whether or not this would alter pathogen tolerance.

Appendix A: Primer sequences

Table A-1. Primer sequences for qPCR.

Gene	AGI	Forward Primer		Reverse Primer		Amplicon Size (cDNA)
		Name	Sequence	Name	Sequence	
<i>AtPQL1</i>	At4g20100	<i>AtPQL1 QL</i>	CTAAGCGTCCTTCTACCAAA	<i>AtPQL1 QR</i>	CTGCTCCTCCAACCTACAAAC	135
<i>AtPQL2</i>	At2g41050	<i>AtPQL2 QL</i>	CGGTTGAAAAACAGAAGAGA	<i>AtPQL2 QR</i>	CCGGTAAACTAGTTCGTTG	170
<i>AtPQL3</i>	At4g36850	<i>AtPQL3 QL</i>	TTGCTTGATGCTATTGTCTG	<i>AtPQL3 QR</i>	CTTGCTTCCACATAGTCTCC	131
<i>AtPQL4</i>	At5g59470	<i>AtPQL4 QL</i>	CTAGGAATCGACTTGAGCTG	<i>AtPQL4 QR</i>	GAAGCTTAACGGTCATTGAA	133
<i>AtPQL5</i>	At5g40670	<i>AtPQL5 QL</i>	CGATTTTACTGGAGGACTTG	<i>AtPQL5 QR</i>	AATCAATGAGTGGTTCGTTT	224
<i>AtPQL6</i>	At4g07390	<i>AtPQL6 QL</i>	GCATAAAGGTCTTCCCTTTT	<i>AtPQL6 QR</i>	GATTAATCTGTCCAGCAAGC	173
<i>YLS8</i>	At5g08290	<i>YLS8 QL</i>	ACTGGGATGAGACCTGTATG	<i>YLS8 QR</i>	TGTTGTTGTTACCAGTTCCA	204
<i>TUB5</i>	At1g20010	<i>TUB5 QL</i>	TCAACAAATGTGGGACTCCC	<i>TUB5 QR</i>	GGTGGGATATCACAACGC	204
<i>BiP3</i>	At1g09080	<i>BiP3 QL</i>	CACGGTTCAGCGTATTTCAT	<i>BiP3 QR</i>	ATAAGCTATGGCAGCACCCGTT	117

Table A-2. Primer sequences for Gateway pENTR-D-TOPO cloning. CACC overhang for TOPO cloning is indicated in red.

Gene	Entry Clone	Forward Primer		Reverse Primer	
		Name	Sequence	Name	Sequence
<i>AtPQL1</i>	<i>pENTR-D-TOPO Pro_{AtPQL1}</i>	<i>AtPQL1 PRO TOPO F</i>	CACCCTCCACGATCAAGAAGCAGAA	<i>AtPQL1 PRO TOPO R</i>	CGTTCTTCACGTATGATACAGC
	<i>pENTR-D-TOPO AtPQL1 NS</i>	<i>AtPQL1 TOPO F</i>	CACCATGATTCGAGATGATTTGTCAC	<i>AtPQL1 TOPO R</i>	GACAGCTTCTTCACCGGTTTCA

Table A-3. Primer sequences for Gateway BP cloning. Partial attB1 (red) and attB2 (blue) recombination sites are indicated.

Gene	Entry Clone	Forward Primer (attB1)		Reverse Primer (attB2)	
		Name	Sequence	Name	Sequence
<i>AtPQL1</i>	<i>pENTR</i> 201 <i>AtPQL1</i>	12attB1- <i>AtPQL1</i>	AAAAAGCAGGCTTCGGTTTAAAC AGGTTTATT	12attB2- <i>AtPQL1</i>	AGAAAGCTGGGTGCCCTAAGTTACATGAA TC
<i>AtPQL2</i>	<i>pENTR</i> 201 <i>AtPQL2</i>	12attB1- <i>AtPQL2</i>	AAAAAGCAGGCTCCAAGATGTTT CTCCACAGC	12attB2- <i>AtPQL2</i>	AGAAAGCTGCGTTTAGACAGCTTCTGCGG T
	<i>pENTR</i> 201 <i>AtPQL2</i>			12attB2- <i>AtPQL2</i> NS	AGAAAGCTGCGTTTGGACAGCTTCTGCGG T
<i>AtPQL3</i>	<i>pENTR</i> 201 <i>AtPQL3</i>	12attB1- <i>AtPQL3</i>	AAAAAGCAGGCTGCCATATGGTG TCTTTAGGT	12attB2- <i>AtPQL3</i>	AGAAAGCTGGGTTTCACGAAACAAAAGTT T
	<i>pENTR</i> 201 <i>AtPQL3</i>			12attB2- <i>AtPQL3</i> NS	AGAAAGCTGGGTGAACAAAAGTTTGTCTT G
<i>AtPQL4</i>	<i>pENTR</i> 201 <i>AtPQL4</i>	12attB1- <i>AtPQL4</i>	AAAAAGCAGGCTTCACCATGGAT TATCTAGGAA	12attB2- <i>AtPQL4</i>	AGAAAGCTGGGTATATCATGAAATCTTTTT AC
	<i>pENTR</i> 201 <i>AtPQL4</i>			12attB2- <i>AtPQL4</i> NS	AGAAAGCTGGGTGTGAAATCTTTTACTTT T
<i>AtPQL5</i>	<i>pENTR</i> 201 <i>AtPQL5</i>	12attB1- <i>AtPQL5</i>	AAAAAGCAGGCTTATCGAGATCG AGGAAAAT	12attB2- <i>AtPQL5</i>	AGAAAGCTGGGTTTAAACATGTTTCATGA GAAG
	<i>pENTR</i> 201 <i>AtPQL5</i>			12attB2- <i>AtPQL5</i> NS	AGAAAGCTGGGTAAACATGTTTCATGAGAA GA
<i>AtPQL6</i>	<i>pENTR</i> 201 <i>AtPQL6</i>	12attB1- <i>AtPQL6</i>	AAAAAGCAGGCTTATCAATGGATT ATCTCG	12attB2- <i>AtPQL6</i>	AGAAAGCTGGGTATCAATTGGCTTTCTTC T
	<i>pENTR</i> 201 <i>AtPQL6</i>			12attB2- <i>AtPQL6</i> NS	AGAAAGCTGGGTGATTGGCTTTCTTCTCC TT
n/a	all	attB1	GGGGACAAGTTTGTACAAAAAAG CAGGCT	attB2	GGGGACCACTTTGTACAAGAAAGCTGGG T

Table A-4. Primers for cloning *AtPQL* open reading frames into pDR195 vectors. NotI (red) and BamHI (blue) overhangs are indicated.

Gene	Plasmid	Forward Primer (NotI)		Reverse Primer (BamHI)	
		Name	Sequence	Name	Sequence
<i>AtPQL1</i>	<i>pDR195</i> <i>P_{PMA1}::AtPQL1</i>	<i>NotI</i> <i>AtPQL1</i>	GTGCGGCCGCATGATTCGAGATGATT TGTC	<i>BamHI</i> <i>AtPQL1</i>	CAGGATCCTTAGACAGCTTCTTCA CCGG
<i>AtPQL2</i>	<i>pDR195</i> <i>P_{PMA1}::AtPQL2</i>	<i>NotI</i> <i>AtPQL2</i>	GTGCGGCCGCATGTTTCTCCACAGCA GTTT	<i>BamHI</i> <i>AtPQL2</i>	CAGGATCCTTAGACAGCTTCTGCG GGTTT
<i>AtPQL4</i>	<i>pDR195</i> <i>P_{PMA1}::AtPQL4</i>	<i>NotI</i> <i>AtPQL4</i>	GTGCGGCCGCATGGATTATCTAGGAA TCGA	<i>BamHI</i> <i>AtPQL4</i>	CAGGATCCTGAAATCTTTTACTT TTCA
<i>AtPQL5</i>	<i>pDR195</i> <i>P_{PMA1}::AtPQL5</i>	<i>NotI</i> <i>AtPQL5</i>	GTGCGGCCGCATGGCGTCGTGGAAT TCGAT	<i>BamHI</i> <i>AtPQL5</i>	CAGGATCCTTAAACATGTTTCATGA GAAG
<i>AtPQL6</i>	<i>pDR195</i> <i>P_{PMA1}::AtPQL6</i>	<i>NotI</i> <i>AtPQL6</i>	GTGCGGCCGCATGGATTATCTCGGAA TCGA	<i>BamHI</i> <i>AtPQL6</i>	CAGGATCCTCAATTGGCTTTCTTC TCCT

Table A-5. Primers for T-DNA insertion line zygosity testing.

Gene	Line	SALK Acc. No.	Forward Primer		Reverse Primer	
			Name	Sequence	Name	Sequence
<i>AtPQL1</i>	<i>pql1-1</i>	108796	108796F	TCCTGAATCAACTAGCCATG G	108796R	GATAAACATTGCGTGATATAA CAAG
<i>AtPQL3</i>	<i>pql3-1</i>	044346	044346F	GGAGAAGAAGACATGCGTGA G	044346R	ACAGAGTATGCAGGGGAAAGC
<i>AtPQL4</i>	<i>pql4-1</i>	018034	018034F	TTAGGTTGGGACCTCACACA G	018034R	CTGCTTCAATGACCGTTAAG C
	<i>pql4-2</i>	000090	000090F	TCTTCGCCATTATAGCTTTCC	000090R	CAAAGCATGCTAAGTGGAG C
<i>AtPQL5</i>	<i>pql5-1</i>	118698	118698F	CATTTGCTGTTTCTGTGATAA TTG	118698R	TTCTTAGGCTTGCCTTTGTCT C
<i>AtPQL6</i>	<i>pql6-1</i>	151293	151293F	TCAGAATCAAGTCGAAGCCA G	151293R	TTAGGCTCTGTGTGTGTGTG TG
	<i>pql6-2</i>	027769	027769F	TGTGTCGACAAATAAACAAAA CTG	027769R	AACTTCAAGCTCAAATGCCAC
-	-	-	Lba1	TGGTTCACGTAGTGGGCCAT CG	all RP	-

Table A-6. Other primers used in this study.

Forward Primer		Reverse Primer	
Name	Sequence	Name	Sequence
ACT2A	CTTACAATTTCCCGCTCTGC	ACT2A	GTTGGGATGAACCAGAAGGA

Appendix B: Supplementary microarray data

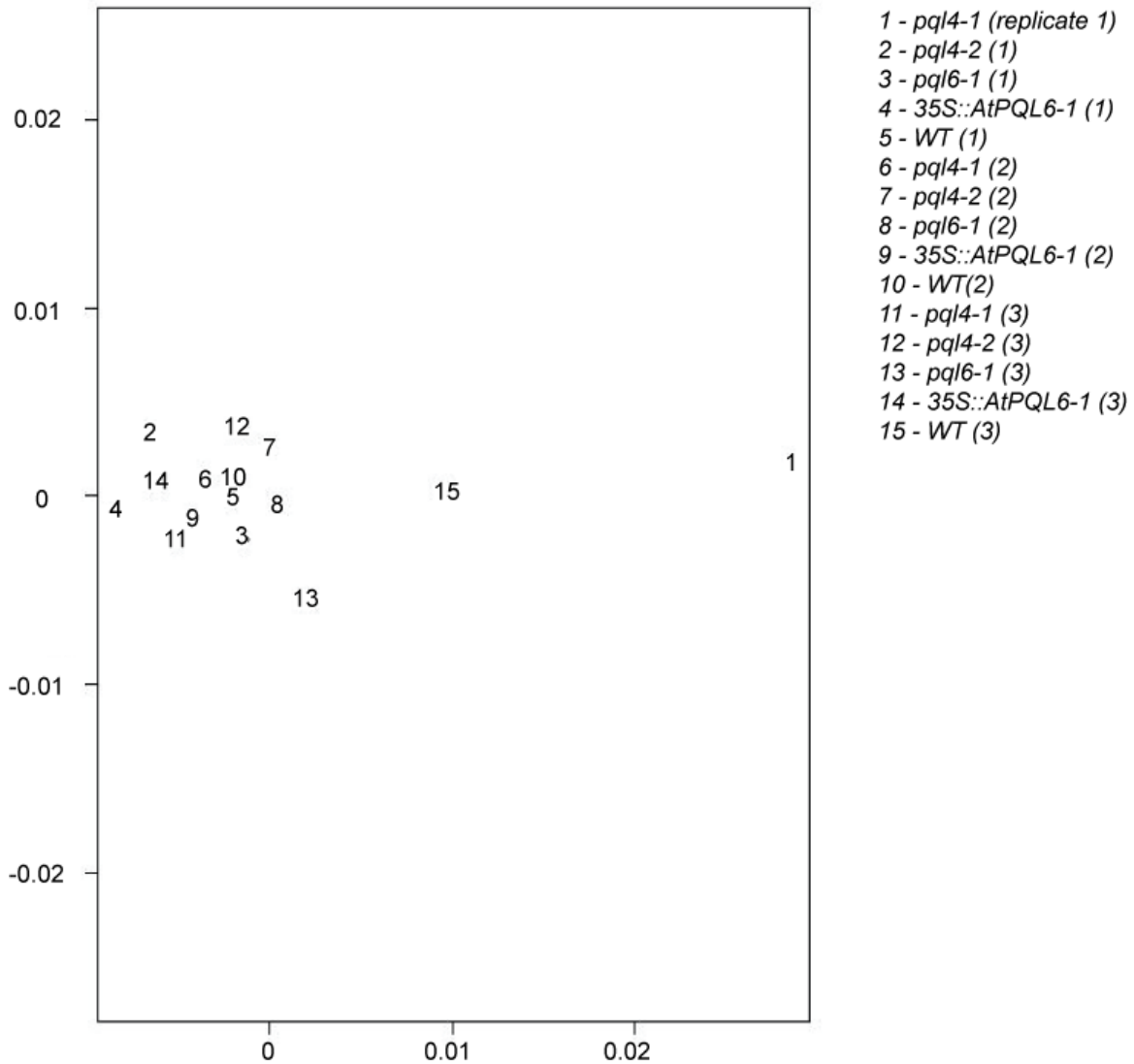


Figure B-1. Sammon mapping of microarray data. In each case the pairwise distance between individual microarray samples in an N-dimensional space (where N is equal to the number of genes) is calculated based on a Pearson correlation between samples. Using the Sammon mapping algorithm the points relating to each sample are then represented in a 2-dimensional plot in a way that best represents their distribution in an N-dimensional space. The closer two points are in the Sammon map the more similar they are.

Table B-1. Genes up-regulated in *pql4-1* relative to wild type. Genes are sorted by false detection rate (FDR) as determined by RP analysis. All genes with an FDR below 10% are shown. For each protein encoded by differentially regulated genes the sub-cellular localisation and number of N-glycosylation sites were predicted. Sub-cellular localisation was predicted using TargetP (c = chloroplast, m = mitochondrion, s = secretory pathway, u = undefined, o = other). The number of N-glycosylation sites was predicted using NetNGlyc (see materials and methods). Proteins predicted to be GPI anchored by Eisenhaber et al. (2003) are also indicated.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glycosylation sites	GPI anchoring
1	264400_at	At1g61800	37.05	1	2.29	---	glucose-6-phosphate/phosphate translocator, putative	c		
2	251928_at	At3g53980	94.99	4	1.91	---	seed storage/lipid transfer protein (LTP) family protein	s		

Table B-2. Genes up-regulated in *pql4-2* relative to wild type. Annotation as for B-1.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glycosylation sites	GPI anchoring
1	244985_at	rpl22	48.61	1	2.41	---	unknown	c		
4	248807_at	At5g47500	105.41	2	2.03	---	pectinesterase family protein	s	4	
2	247717_at	At5g59320	78.99	2	1.95	<i>LTP3</i>	lipid transfer protein 3 (LTP3)	s		
5	249082_at	At5g44120	110.81	2.2	3.03	<i>CRA1</i>	12S seed storage protein (CRA1)	u		
3	251928_at	At3g53980	102.86	2.67	1.9	---	seed storage/lipid transfer protein (LTP) family protein	s		
6	253904_at	At4g27140	136.44	4.67	3.01	---	2S seed storage protein 1 / 2S albumin storage protein	s	1	
7	247718_at	At5g59310	147.43	5.14	1.84	<i>LTP4</i>	lipid transfer protein 4 (LTP4)	s		
8	253894_at	At4g27150	164.62	7	3.01	---	2S seed storage protein 2 / 2S albumin storage protein	s	1	

Table B-3. Genes down-regulated in *pql4-2* relative to wild type. Annotation as for B-1.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glycosylation sites	GPI anchoring
3	251428_at	At3g60140	27.52	0	-1.9	---	glycosyl hydrolase family 1 protein	s	2	
1	245693_at	At5g04260	1.22	0	-5.37	---	thioredoxin family protein	c		
2	247780_at	At5g58770	16.66	0	-2.09	<i>DPS-Like</i>	dehydrodolichyl diphosphate (DEDOL-PP) synthase, putative	u		
5	265665_at	At2g27420	62.11	0.8	-1.71	---	cysteine proteinase, putative	s	1	
6	249774_at	At5g24150	65.45	0.83	-1.8	<i>SQP1,1</i>	squalene monooxygenase 1,1	s		
4	256527_at	At1g66100	52.95	1	-1.82	---	thionin, putative	s		
9	244972_at	psbB	99.47	1.44	-1.6	---	unknown	c		
8	263174_at	At1g54040	95.88	1.62	-1.64	---	kelch repeat-containing protein	u		
7	257100_at	At3g25010	94.93	1.86	-1.61	---	disease resistance family protein	s	16	
10	263539_at	At2g24850	108.46	2	-1.73	---	aminotransferase, putative	u		
13	264163_at	At1g65445	125.79	2.31	-1.59	---	transferase-related	u		
12	266292_at	At2g29350	123.65	2.33	-1.59	---	tropinone reductase, putative	u		
11	254975_at	At4g10500	117.55	2.45	-1.62	---	2OG-Fe(II) oxygenase family protein	u		
14	249732_at	At5g24420	143.05	2.79	-1.55	---	glucosamine/galactosamine-6-phosphate isomerase-related	o		
15	261459_at	At1g21100	152.93	3	-1.55	---	O-methyltransferase, putative	o		
16	262412_at	At1g34760	156.95	3	-1.55	<i>GRF11</i>	14-3-3 protein GF14 omicron	o		
17	253046_at	At4g37370	189.88	5.06	-1.55	---	cytochrome P450, putative	u	3	
18	248169_at	At5g54610	195.69	5.17	-1.49	---	ankyrin repeat family protein	u		
23	245265_at	At4g14400	233.07	6.61	-1.48	---	ankyrin repeat family protein	u		
19	245007_at	psaA	217.75	6.74	-1.48	---	unknown	c		
25	267516_at	At2g30520	244.06	6.8	-1.5	<i>RPT2</i>	signal transducer of phototropic response	u		
24	264436_at	At1g10370	239.8	6.88	-1.53	<i>ERD9</i>	glutathione S-transferase, putative	u		
22	261059_at	At1g01250	232.99	6.91	-1.51	---	AP2 domain-containing transcription factor	c		
21	256446_at	At3g11110	231.66	7.14	-1.5	---	zinc finger family	o		
20	261339_at	At1g35710	227.17	7.25	-1.49	---	LRR protein kinase, putative	s	12	
28	249661_at	At1g55265	263.19	7.68	-1.49	---	expressed protein	s		
27	247949_at	At5g57220	258.85	7.78	-1.46	---	cytochrome P450, putative	s		
26	267344_at	At2g44230	257.61	7.92	-1.47	---	expressed protein	c		
29	256548_at	At3g14770	281.77	8.79	-1.49	---	nodulin MtN3 family protein	s		
30	245928_s_at	At5g24780	287.69	9	-1.52	<i>VSP2</i>	vegetative storage protein 2	s	1	
31	252414_at	At3g47420	297.37	9.39	-1.49	---	glycerol-3-phosphate transporter, putative	o		
32	267038_at	At2g38480	305.3	9.75	-1.49	---	integral membrane protein, putative	o		

Table B-4. Genes down-regulated in *pql6-1* relative to wild type. Annotation as for B-1.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glycosylation sites	GPI anchoring
1	255199_at	At4g07390	1.51	0	-5.23	<i>AtPQL6</i>	PQ-loop repeat family protein	o		
3	256503_at	At1g75250	74.09	1	-1.78	---	myb family transcription factor	u		
2	249477_s_at	At5g38930	66.86	1	-2.31	---	germin-like protein, putative	s	1	
4	249645_at	At5g36910	98.12	2.25	-1.62	<i>THI2.2</i>	thionin	s	1	
5	266383_at	At2g14580	123.81	4.4	-1.76	---	pathogenesis-related protein, putative	s		
7	265478_at	At2g15890	179.04	9	-1.59	<i>MEE14</i>	expressed protein	u		

Table B-5. Genes up-regulated in *pql6-1* relative to wild type. Annotation as for B-1.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
1	264400_at	At1g61800	7.89	0	3.05	---	glucose-6-phosphate translocator, putative	c		
3	266385_at	At2g14610	37.27	0	2.5	<i>PR-1</i>	pathogenesis-related protein 1 (PR-1)	s	1	
2	246302_at	At3g51860	17.86	0	2.48	---	cation exchanger, putative (CAX3)	u		
11	252511_at	At3g46280	72.47	0.45	1.95	---	protein kinase-related	s	3	
12	263536_at	At2g25000	79.2	0.5	1.9	---	WRKY family transcription factor	o		
10	250302_at	At5g11920	69.77	0.5	1.99	---	glycosyl hydrolase family 32 protein	u		
15	254975_at	At4g10500	93.41	0.53	1.85	---	2OG-Fe(II) oxygenase family protein	u		
13	251625_at	At3g57260	85.55	0.54	1.89	<i>PR-2</i>	glycosyl hydrolase family 17 protein	s		
9	263539_at	At2g24850	68.9	0.56	2.09	---	aminotransferase, putative	u		
14	254805_at	At4g12480	91.21	0.57	1.94	---	lipid transfer protein (LTP) family protein	s		
8	255595_at	At4g01700	68.43	0.62	1.94	---	chitinase, putative	s	1	
7	262050_at	At1g80130	66.66	0.71	1.97	---	expressed protein	u		
17	254818_at	At4g12470	101.8	0.71	1.88	---	lipid transfer protein (LTP) family protein	s		
18	249454_at	At5g39520	108.51	0.72	1.85	---	expressed protein	c		
19	259507_at	At1g43910	111.87	0.74	1.75	---	AAA-type ATPase family protein	s		
16	254234_at	At4g23680	101.09	0.75	1.92	---	major latex protein-related / MLP-related	o		
5	262482_at	At1g17020	63.19	0.8	1.97	---	2OG-Fe(II) oxygenase family protein	u		
6	258941_at	At3g09940	64.6	0.83	1.98	---	monodehydroascorbate reductase	u		
20	264648_at	At1g09080	124.19	0.9	1.77	<i>BiP-3</i>	luminal binding protein 3 (BiP-3)	s		
4	260919_at	At1g21520	56.04	1	2.02	---	expressed protein	u		
21	260140_at	At1g66390	129.75	1.1	1.83	<i>PAP2</i>	myb family transcription factor, putative	o		
28	259550_at	At1g35230	151.78	1.54	1.71	<i>AGP5</i>	arabinogalactan-protein	u		Yes
22	266368_at	At2g41380	140.93	1.59	1.74	---	embryo-abundant protein-related	u		
27	247800_at	At5g58570	151.41	1.59	1.81	---	phosphatidate cytidyltransferase family	o		
25	254042_at	At4g25810	147.73	1.6	1.72	<i>XTR6</i>	xyloglucan:xyloglucosyl transferase	s	2	
26	257365_x_at	At2g26020	150.36	1.62	1.98	<i>PDF1.2b</i>	plant defensin-fusion protein	s		
24	252131_at	At3g50930	145.84	1.67	1.71	---	AAA-type ATPase family protein	u		
23	249052_at	At5g44420	145.79	1.74	1.95	<i>PDF1.2a</i>	plant defensin protein, putative	s		
29	248236_at	At5g53870	160.4	1.86	1.74	---	plastocyanin-like domain protein	s	2	Yes
30	264635_at	At1g65500	163.42	1.9	1.79	---	expressed protein	s		
31	252170_at	At3g50480	165.79	1.97	1.87	---	broad-spectrum mildew resistance family	s		
32	254265_s_at	At4g23140	181.8	2.47	1.72	<i>RLK5</i>	receptor-like protein kinase 5	s	9	
33	256245_at	At3g12580	185.97	2.58	1.54	---	heat shock protein 70, putative	o		
34	248169_at	At5g54610	196.75	2.71	1.63	---	ankyrin repeat family protein	u		
42	265161_at	At1g30900	219.46	2.93	1.74	<i>AtELP6</i>	vacuolar sorting receptor, putative	s	3	
41	261545_at	At1g63530	218.36	2.98	1.59	---	hypothetical protein	o		
43	246368_at	At1g51890	223.81	3.02	1.73	---	leucine-rich repeat protein kinase, putative	s	17	
40	254819_at	At4g12500	217.47	3.02	1.75	---	lipid transfer protein (LTP) family protein	u	1	
39	254832_at	At4g12490	216.49	3.08	1.77	---	lipid transfer protein (LTP) family protein	u		
35	252549_at	At3g45860	207.96	3.14	1.6	---	receptor-like protein kinase, putative	s	9	
38	250286_at	At5g13320	215.4	3.16	1.62	---	auxin-responsive GH3 family protein	o		

Table B-5 continued. Genes up-regulated in *pql6-1* relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
37	260556_at	At2g43620	213.69	3.22	1.73	---	chitinase, putative	s		
36	252136_at	At3g50770	211.16	3.22	1.66	---	calmodulin-related protein, putative	c		
44	266292_at	At2g29350	244.13	3.75	1.57	---	tropinone reductase, putative	u		
45	262118_at	At1g02850	245.63	3.76	1.59	---	Glycosyl hydrolase family 1 protein	s	3	
46	263584_at	At2g17040	246.55	3.76	1.73	---	no apical meristem (NAM) family protein	o		
48	252417_at	At3g47480	257.3	4.23	1.79	---	calcium-binding EF hand family protein	s		
49	266743_at	At2g02990	260.45	4.27	1.61	---	ribonuclease 1 (RNS1)	s		
47	266782_at	At2g29120	257.07	4.32	1.57	<i>GLR2.7</i>	glutamate receptor family protein	s	10	
50	250445_at	At5g10760	264.94	4.32	1.56	---	aspartyl protease family protein	s	2	
53	247327_at	At5g64120	271.72	4.4	1.66	---	peroxidase, putative	s	1	
52	261240_at	At1g32940	269.19	4.46	1.66	---	subtilase family protein	s	7	
54	246620_at	At5g36220	276.05	4.46	1.54	<i>CYP81D1</i>	cytochrome P450 81D1 (CYP81D1)	s	1	
51	257763_s_at	At3g23110	267.83	4.47	1.58	---	Disease resistance family protein	s	13	
55	250435_at	At5g10380	279.57	4.64	1.58	---	zinc finger family protein	u		
56	245228_at	At3g29810	289.98	4.98	1.53	<i>COBL2</i>	COBRA cell expansion protein	s	9	Yes
57	247604_at	At5g60950	296.98	5.35	1.67	---	phytochelatin synthetase-related	s	2	
61	260568_at	At2g43570	319.01	5.98	1.58	---	chitinase, putative	s	1	
60	252888_at	At4g39210	317.82	6.02	1.65	<i>APL3</i>	glucose-1-phosphate adenyltransferase	o		
59	255630_at	At4g00700	317.62	6.12	1.5	---	C2 domain-containing protein	u		
58	255406_at	At4g03450	314.96	6.12	1.59	---	ankyrin repeat family protein	u		
62	263783_at	At2g46400	336.89	6.74	1.52	---	WRKY family transcription factor	u		
63	246366_at	At1g51850	344.82	6.95	1.64	---	leucine-rich repeat kinase, putative	s	12	
64	252068_at	At3g51440	348.29	7.12	1.48	---	strictosidine synthase family protein	s	2	
66	264777_at	At1g08630	354.25	7.27	1.44	---	L-allo-threonine aldolase-related	u		
65	245325_at	At4g14130	353.11	7.32	1.34	<i>XTR7</i>	xyloglucan endotransglycosylase, putative	s	1	
67	245306_at	At4g14690	365.46	7.78	1.54	---	chlorophyll A-B binding family	u		
70	254314_at	At4g22470	373.36	7.79	1.6	---	lipid transfer protein (LTP) family protein	s	3	
69	256017_at	At1g19180	370.36	7.8	1.49	---	expressed protein	u		
68	262832_s_at	At1g14870	367.5	7.84	1.62	---	expressed protein	o		
72	246099_at	At5g20230	376.78	7.94	1.56	---	plastocyanin-like domain protein	s	1	Yes
71	267168_at	At2g37770	376.06	7.97	1.5	---	aldo/keto reductase family protein	u		
73	254833_s_at	At4g12280	380.5	8.11	1.54	---	Copper amine oxidase, putative	u		
74	249917_at	At5g22460	384.97	8.3	1.49	---	esterase/lipase/thioesterase family protein	s		
80	262703_at	At1g16510	416.68	9.64	1.48	---	auxin-responsive family protein	u		
82	267546_at	At2g32680	421.6	9.73	1.46	---	Disease resistance family protein	s	16	
79	255243_at	At4g05590	416.37	9.75	1.67	---	expressed protein	u		
81	265214_at	At1g05000	420.96	9.81	1.43	---	tyrosine specific phosphatase family	o		
78	256169_at	At1g51800	416.31	9.87	1.56	---	leucine-rich repeat kinase, putative	s	15	
76	257315_at	At3g30775	413.32	9.87	1.48	<i>POX</i>	proline dehydrogenase	m		
75	254574_at	At4g19430	412.69	9.95	1.43	---	expressed protein	c		
77	253181_at	At4g35180	415.68	9.96	1.51	---	amino acid transporter family protein	u		

Table B-6. Genes up-regulated in *35S::AtPQL6-1* relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
2	245016_at	accD	10.71	0	3.01	---	Unknown	c		
1	255199_at	At4g07390	6.06	0	4.01	<i>AtPQL6</i>	PQ-loop repeat family protein	o		
5	249863_at	At5g22950	33.76	0	2.44	---	SNF7 family protein	u		
4	249271_at	At5g41790	25.76	0	2.63	<i>CIP1</i>	COP1-interactive protein 1	u		
3	244985_at	rpl22	17	0	3.2	---	Unknown	c		
7	262958_at	At1g54410	50.98	0.14	2.47	---	dehydrin family protein	o		
6	247600_at	At5g60890	49.06	0.17	2.43	<i>ATR1</i>	receptor-like protein kinase (ATR1) (MYB34)	o		
9	260031_at	At1g68790	67.2	0.22	2.21	---	expressed protein	u		
8	259415_at	At1g02330	65.99	0.25	2.22	---	expressed protein	u		
13	256744_at	At3g29350	78.87	0.38	2.18	<i>HP1</i>	two-component phosphorelay mediator 1 (HP1)	u		
10	262871_at	At1g65010	70.72	0.4	2.18	---	expressed protein	c		
12	255588_at	At4g01570	77.02	0.42	2.07	---	pentatricopeptide (PPR) repeat-containing protein	u		
14	250919_at	At5g03660	80.13	0.43	2.25	---	expressed protein	u		
16	246396_at	At1g58180	82.85	0.44	2.2	---	carbonic anhydrase family protein	m		
11	248106_at	At5g55100	75.78	0.45	2.15	---	SWAP domain-containing protein	u		
15	247942_at	At5g57120	80.81	0.47	2.26	---	expressed protein	u		
17	251272_at	At3g61890	91.6	0.53	2.08	<i>ATHB-12</i>	homeobox-leucine zipper protein 12 (HB-12)	o		
25	244980_at	rpl36	99.42	0.56	2.27	---	Unknown	c		
23	258017_at	At3g19370	97.09	0.57	2.03	---	expressed protein	u		
24	251887_at	At3g54170	98.45	0.58	2.26	<i>FIP37</i>	FKBP12 interacting protein (FIP37)	o		
22	262204_at	At2g01100	96.95	0.59	2.13	---	expressed protein	o		
21	250863_at	At5g04750	95.53	0.62	2.25	---	F1F0-ATPase inhibitor protein, putative	u		
27	244981_at	rps8	105.46	0.63	2.31	---	Unknown	c		
26	244982_at	rpl14	103.24	0.65	2.28	---	Unknown	c		
20	244984_at	rps3	95.26	0.65	2.27	---	Unknown	c		
18	247696_at	At5g59780	94.25	0.67	2.2	---	myb family transcription factor (MYB59)	u		
19	247977_at	At5g56850	94.92	0.68	2.22	---	expressed protein	u		
29	259642_at	At1g69030	113.39	0.69	2.31	---	BSD domain-containing protein	u		
28	264648_at	At1g09080	112.55	0.71	1.95	<i>BiP-3</i>	luminal binding protein 3 (BiP-3)	s		
30	247252_at	At5g64770	116.1	0.73	2.36	---	expressed protein	s	1	
31	244983_at	rpl16	119.57	0.77	2.47	---	Unknown	c		
32	258245_at	At3g29075	122.08	0.81	2.01	---	glycine-rich protein	o		
42	261031_at	At1g17360	149.94	0.93	1.9	---	COP1-interacting protein-related	u		
41	260551_at	At2g43510	149.85	0.95	2.03	---	trypsin inhibitor, putative	s		
34	263112_at	At1g03080	137.88	0.97	1.92	---	kinase interacting family protein	u		
40	253294_at	At4g33750	149.57	0.97	1.94	---	expressed protein	u		
44	266537_at	At2g16860	151.91	0.98	2.11	---	GCIP-interacting family protein	u		
43	249322_at	At5g40930	150.97	0.98	2.05	<i>TOM20-4</i>	mitochondrial import receptor subunit TOM20-4	o		
45	248291_at	At5g53020	153.62	0.98	1.85	---	expressed protein	o		
33	261447_at	At1g21160	136.51	1	1.97	<i>ELF2</i>	eukaryotic translation initiation factor 2 family protein	u		
47	262984_at	At1g54460	162.2	1	2.01	---	expressed protein	o		
39	266322_at	At2g46690	149.03	1	1.97	---	auxin-responsive family protein	u		
38	255171_at	At4g07990	148.23	1	2.02	---	DNAJ heat shock N-terminal domain-containing protein	u		
35	249340_at	At5g41140	140.51	1	2	---	expressed protein	u		
36	248826_at	At5g47080	144.72	1	2.05	---	casein kinase II beta chain, putative	u		
46	255795_at	At2g33380	161.57	1.02	1.8	---	calcium-binding RD20 protein (RD20)	u		
49	249441_at	At5g39730	166.76	1.02	2.2	---	avirulence induced gene (AIG) protein-related	u		
37	248322_at	At5g52760	147.42	1.03	1.98	---	heavy-metal-associated domain-containing protein	u		
48	253359_at	At4g33080	166.54	1.04	1.91	---	protein kinase, putative	o		
50	248109_at	At5g55310	169.4	1.08	2	---	DNA topoisomerase I, putative	o		
51	255437_at	At4g03060	176.95	1.2	1.87	<i>AOP2</i>	2-oxoglutarate-dependent dioxygenase, putative	o		
53	258126_at	At3g24490	181.64	1.28	1.97	---	expressed protein	o		
52	254234_at	At4g23680	181.4	1.29	1.88	---	major latex protein-related / MLP-related	o		
55	263296_at	At2g38800	193.39	1.49	1.85	---	calmodulin-binding protein-related	u		
56	246493_at	At5g16180	194.19	1.5	1.93	---	hypothetical protein	u		
54	248146_at	At5g54940	193.32	1.52	1.91	---	eukaryotic translation initiation factor SU11, putative	u		
57	255243_at	At4g05590	195.24	1.54	1.94	---	expressed protein	u		
58	256379_at	At1g66840	198.76	1.64	1.81	---	expressed protein	u		
59	246453_at	At5g16830	209.55	1.88	1.83	<i>SYP21</i>	syntaxin 21 / PEP12 homolog	u		

Table B-6 continued. Genes up-regulated in *35S::AtPQL6-1* relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
61	253977_at	At4g26630	217.57	1.97	1.85	---	expressed protein	o		
68	260012_at	At1g67865	233.43	2.24	1.95	---	expressed protein	u		
63	263220_at	At1g30610	228.19	2.27	1.81	---	pentatricopeptide (PPR) repeat-containing protein	u		
64	258081_at	At3g26085	229.65	2.27	1.83	---	CAAX amino terminal protease family protein	u		
67	247002_at	At5g67320	233.28	2.27	1.81	---	WD-40 repeat family protein	s		
66	249425_at	At5g39790	232.48	2.29	1.77	---	5'-AMP-activated protein kinase-related	u		
62	255128_at	At4g08310	228.1	2.31	1.93	---	expressed protein	o		
65	258935_at	At3g10120	232.11	2.32	1.88	---	expressed protein	o		
69	262747_at	At1g28540	240.27	2.38	1.75	---	expressed protein	o		
70	260981_at	At1g53460	242.72	2.43	1.87	---	expressed protein	u		
71	266402_at	At2g38780	245.1	2.51	1.88	---	expressed protein	c		
77	257759_at	At3g23070	253.35	2.52	1.81	---	expressed protein	u		
94	260251_at	At1g74250	274.42	2.54	1.74	---	DNAJ heat shock protein	u		
76	258333_at	At3g16000	252.72	2.54	1.82	---	matrix-localized MAR DNA-binding protein-related	c		
92	256300_at	At1g69490	271.58	2.55	1.73	---	no apical meristem (NAM) family protein	o		
93	252676_at	At3g44280	272.58	2.55	1.9	---	expressed protein	s	2	
91	249409_at	At5g40340	271.1	2.55	1.78	---	PWWP domain-containing protein	o		
90	260446_at	At1g72420	270.54	2.57	1.78	---	chaperone-related	u		
75	266385_at	At2g14610	252.43	2.57	1.94	<i>PR-1</i>	pathogenesis-related protein 1 (PR-1)	s	1	
96	249886_at	At5g22320	278.73	2.58	1.79	---	leucine-rich repeat family protein	o		
88	254269_at	At4g23050	268.85	2.59	1.69	---	protein kinase, putative	u		
74	250179_at	At5g14440	251.81	2.59	1.9	---	surfeit locus 2 protein-related	o		
98	248099_at	At5g55300	280.14	2.59	1.88	---	DNA topoisomerase I	o		
89	252820_at	At3g42640	270.37	2.6	1.8	---	ATPase, plasma membrane-type, putative	o		
73	248344_at	At5g52280	250.91	2.6	1.74	---	protein transport protein-related	u		
78	247331_at	At5g63530	256.63	2.6	1.81	---	copper chaperone (CCH)-related	o		
97	262137_at	At1g77920	279.31	2.61	1.86	---	bZIP family transcription factor	u		
95	255338_at	At4g04470	278.32	2.61	1.76	---	peroxisomal membrane protein 22 kDa (PMP22)	u		
87	261584_at	At1g01040	268.63	2.62	1.77	---	DEAD/DEAH box helicase carpel factory / CAF	o		
72	264507_at	At1g09415	249.78	2.62	1.84	<i>NIMIN-3</i>	NPR1/NIM1-interacting protein 3 (NIMIN-3)	o		
101	250704_at	At5g06265	286.03	2.62	1.82	---	hyaluronan mediated motility receptor-related	s	1	
99	260943_at	At1g45145	284.73	2.64	1.92	---	thioredoxin H-type 5 (TRX-H-5) (TOUL)	o		
100	258507_at	At3g06500	285.97	2.65	1.7	---	beta-fructofuranosidase, putative	u		
86	248186_at	At5g53880	268.6	2.65	2	---	expressed protein	u		
85	251221_at	At3g62550	268.24	2.66	1.93	---	universal stress protein (USP) family protein	o		
83	244937_at	ndhH	266.42	2.67	1.87	---	Unknown	c		
84	252904_at	At4g39620	268.16	2.69	1.87	---	pentatricopeptide (PPR) repeat-containing protein	c		
79	250946_at	At5g03420	263.4	2.7	1.72	---	dentin sialophosphoprotein-related	u		
81	250150_at	At5g14710	265.69	2.7	1.81	---	expressed protein	o		
82	250066_at	At5g17930	266.21	2.71	1.78	---	MA3 domain-containing protein	u		
80	255093_s_at	At4g08580	265.66	2.74	1.76	---	microfibrillar-associated protein-related	u		
102	260824_at	At1g06720	289.47	2.76	1.76	---	expressed protein	u		
103	249612_at	At5g37290	290.95	2.76	1.82	---	armadillo/beta-catenin repeat family protein	u		
104	252188_at	At3g50860	292.93	2.81	1.78	---	clathrin adaptor complex small chain family protein	u		
105	263007_at	At1g54260	295.77	2.95	1.71	---	histone H1/H5 family protein	u		
106	261410_at	At1g07610	297.34	2.96	2.01	<i>MT-1C</i>	metallothionein-like protein 1C (MT-1C)	u		
107	256596_at	At3g28540	301.53	3	1.78	---	AAA-type ATPase family protein	u	3	
108	263331_at	At2g15270	303.14	3.03	1.81	---	expressed protein	u		
113	252433_at	At3g47560	308.41	3.06	1.79	---	esterase/lipase/thioesterase family protein	u		
112	249101_at	At5g43580	307.45	3.06	1.95	---	protease inhibitor, putative	u		
115	252206_at	At3g50360	311.8	3.07	1.9	---	caltractin / centrin	u		
111	254975_at	At4g10500	307.41	3.08	1.6	---	2OG-Fe(II) oxygenase family protein	u		
116	254955_at	At4g10920	313.33	3.09	1.92	---	KELP family protein	o		
110	255895_at	At1g17990	306.99	3.1	1.82	<i>OPR</i>	12-oxophytodienoate reductase, putative	u		
114	252612_at	At3g45160	311.73	3.1	1.97	---	expressed protein	s		
109	253897_at	At4g27120	306.76	3.12	1.8	---	expressed protein	u	1	
117	261845_at	At1g15960	318.17	3.21	1.71	<i>NRAMP6</i>	NRAMP metal ion transporter 6, putative	u		
128	248386_at	At5g51940	330.43	3.22	1.78	---	DNA-directed RNA polymerase II, putative	o		
118	262314_at	At1g70810	320.38	3.24	1.85	---	C2 domain-containing protein	u		
119	248657_at	At5g48570	321.73	3.24	1.59	---	peptidyl-prolyl cis-trans isomerase, putative	o		

Table B-6 continued. Genes up-regulated in *35S::AtPQL6-1* relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
129	248537_at	At5g50100	331.8	3.24	1.83	---	expressed protein	m		
127	248070_at	At5g55660	330.24	3.24	1.68	---	expressed protein	o		
125	265336_at	At2g18290	329.07	3.26	1.79	---	anaphase-promoting complex, subunit 10 family	o		
126	249528_at	At5g38720	330.23	3.26	1.85	---	expressed protein	o		
124	265707_at	At2g03390	328.55	3.27	1.75	---	uvrB/uvrC motif-containing protein	c		
130	247032_at	At5g67240	333.17	3.28	1.8	---	exonuclease family protein	o		
123	259794_at	At1g64330	327.78	3.29	1.74	---	myosin heavy chain-related	o		
122	247115_at	At5g65930	326.44	3.3	1.81	<i>ZWICHEL</i>	kinesin-like calmodulin-binding protein	u		
121	265044_at	At1g03950	325.38	3.31	1.71	---	SNF7 family protein	u		
120	251759_at	At3g55630	324.85	3.33	1.69	<i>FPGS4</i>	folylpolyglutamate synthetase	o		
134	257717_at	At3g18390	343.29	3.45	1.82	---	expressed protein	c		
133	247903_at	At5g57340	342.89	3.47	1.73	---	expressed protein	u		
131	251309_at	At3g61220	341.81	3.48	1.75	---	short-chain dehydrogenase/reductase (SDR) family	u		
132	247733_at	At5g59610	342.46	3.48	1.87	---	DNAJ heat shock protein	u		
135	262794_at	At1g13120	346.29	3.49	1.79	---	expressed protein	u		
136	249628_at	At5g37580	348.52	3.51	1.71	---	tropomyosin-related	m		
137	263569_at	At2g27170	350.14	3.52	1.61	---	structural maintenance of chromosomes family	u		
139	265511_at	At2g05540	352.02	3.53	1.58	---	glycine-rich protein	s		
138	255444_at	At4g02560	350.94	3.54	1.7	---	homeobox protein LUMINIDEPENDENS (LD)	o		
140	254335_at	At4g22260	357.05	3.61	1.86	<i>IM</i>	alternative oxidase, putative	c		
141	262010_at	At1g35612	366.97	3.89	1.72	---	expressed protein	u		
142	263536_at	At2g25000	369.32	3.92	1.59	---	WRKY family transcription factor	o		
143	263488_at	At2g31840	372.43	3.99	1.72	---	expressed protein	c		
145	248310_at	At5g52550	379.33	4.1	1.81	---	expressed protein	u		
144	249378_at	At5g40450	378.82	4.11	1.67	---	expressed protein	o		
146	256754_at	At3g25690	383.75	4.2	1.73	---	hydroxyproline-rich glycoprotein	s	5	
159	254848_at	At4g11960	396.53	4.22	1.57	---	expressed protein	c		
158	247235_at	At5g64580	396	4.22	1.69	---	AAA-type ATPase family protein	c		
153	261500_at	At1g28400	392.54	4.24	1.74	---	expressed protein	s		
157	258933_at	At3g09980	395.64	4.24	1.85	---	expressed protein	u		
147	262608_at	At1g14120	386.45	4.27	1.45	---	2-oxoglutarate-dependent dioxygenase, putative	u		
156	261117_at	At1g75310	395.42	4.27	1.6	<i>AUL1</i>	DNAJ heat shock protein	u		
152	246936_at	At5g25360	392.54	4.27	1.8	---	expressed protein	u		
149	258736_at	At3g05900	388.94	4.28	1.62	---	neurofilament protein-related	o		
150	259362_s_at	At1g13350	391.25	4.29	1.74	---	protein kinase family protein	u		
148	245280_at	At4g16845	388.4	4.29	1.77	<i>VRN1</i>	vernalization 2 protein (VRN2)	u		
151	260598_at	At1g55930	392.17	4.3	1.62	---	CBS domain-containing protein	c		
155	245302_at	At4g17695	395.31	4.3	1.7	---	myb family transcription factor (KAN3)	u		
154	251057_at	At5g01780	394.55	4.3	1.68	---	oxidoreductase, 2OG-Fe(II) oxygenase family	u		
160	263539_at	At2g24850	402.61	4.39	1.57	---	aminotransferase, putative	u		
161	260116_at	At1g33960	405.07	4.47	1.66	<i>AIG1</i>	avirulence induced gene 1	o		
162	247370_at	At5g63320	408.17	4.52	1.55	---	expressed protein	o		
165	245621_at	At4g14070	411.29	4.53	1.71	---	AMP-binding protein, putative	c		
163	261819_at	At1g11410	410.42	4.55	1.55	---	S-locus protein kinase, putative	s	8	
164	266298_at	At2g29590	411.03	4.55	1.67	---	thioesterase family protein	o		
166	250866_at	At5g03905	413.89	4.58	1.69	---	hesB-like domain-containing protein	u		
167	265678_at	At2g31970	415.08	4.59	1.6	<i>RAD50</i>	DNA repair-recombination protein	o		
168	249285_at	At5g41960	420.01	4.74	1.7	---	expressed protein	c		
169	251799_at	At3g55520	421.62	4.78	1.7	---	peptidyl-prolyl cis-trans isomerase, putative	o		
170	258338_at	At3g16150	422.91	4.79	1.52	---	L-asparagine amidohydrolase, putative	u		
171	260660_at	At1g19485	425.85	4.84	1.6	---	AT hook motif-containing protein	o		
174	246957_at	At5g24670	428.15	4.89	1.65	---	cytidine/deoxycytidylate deaminase family	o		
173	261422_at	At1g18730	427.94	4.91	1.89	---	expressed protein	c		
172	257007_at	At3g14205	427.52	4.93	1.68	---	phosphoinositide phosphatase family	u		
177	266293_at	At2g29360	432.25	4.97	1.68	---	tropine dehydrogenase, putative	u		
176	254412_at	At4g21430	431.83	4.97	1.61	---	jumonji (jmjC) domain-containing	u		
180	266795_at	At2g03070	435.99	4.98	1.53	---	expressed protein	o		
179	267645_at	At2g32860	434.29	4.99	1.74	---	glycosyl hydrolase family 1	s	3	
175	254019_at	At4g26190	431.72	4.99	1.57	---	expressed protein	o		
178	252712_at	At3g43800	434.07	5.02	1.73	---	glutathione S-transferase, putative	u		

Table B-6 continued. Genes up-regulated in 35S::AtPQL6-1 relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
181	247273_at	At5g64340	439.49	5.02	1.65	---	expressed protein	u		
182	261151_at	At1g19650	443.12	5.13	1.7	---	phosphoglyceride transfer protein, putative	u		
183	266572_at	At2g23840	444.7	5.15	1.64	---	HNH endonuclease domain-containing	u		
184	255088_at	At4g09350	448.53	5.26	1.54	---	DNAJ heat shock protein	u		
188	249308_at	At5g41350	452.98	5.26	1.7	---	zinc finger (C3HC4-type RING finger) family	u		
186	246476_at	At5g16730	451.22	5.27	1.66	---	expressed protein	c		
190	262809_at	At1g11720	455.99	5.28	1.64	---	starch synthase, putative	c		
187	264126_at	At1g79280	452.38	5.28	1.63	---	expressed protein	o		
185	257178_at	At3g13070	450.6	5.3	1.7	---	CBS domain-containing	c		
189	267477_at	At2g02710	455.61	5.31	1.55	---	PAC motif-containing	o		
191	249386_at	At5g40060	457.23	5.31	1.67	---	disease resistance protein (TIR-NBS class)	u		
192	254305_at	At4g22200	459.82	5.38	1.49	AKT2/3	potassium channel protein 2 (AKT2/AKT3)	u		
193	259219_at	At3g03560	461.13	5.39	1.65	---	expressed protein	o		
195	264698_at	At1g70200	463.25	5.44	1.86	---	RNA recognition motif (RRM)-containing	c		
194	254083_at	At4g24920	462.79	5.45	1.89	---	transport protein SEC61 gamma subunit, putative	o		
196	251932_at	At3g54010	464.62	5.47	1.66	PAS1-D	peptidyl-prolyl cis-trans isomerise	o		
198	251182_at	At3g62600	471.23	5.64	1.69	---	DNAJ heat shock family protein	s	1	
197	261886_s_at	At1g80700	471.02	5.66	1.73	---	expressed protein	m		
202	261206_at	At1g12800	475.71	5.68	1.77	---	S1 RNA-binding domain-containing protein	c		
201	261133_at	At1g19720	474.62	5.68	1.55	---	jacalin lectin family	u		
200	265698_at	At2g32160	473.85	5.68	1.67	---	expressed protein	u		
199	249652_at	At5g37070	473.59	5.71	1.64	---	expressed protein	u		
204	247854_at	At5g58200	477.49	5.72	1.65	---	expressed protein	o		
205	253344_at	At4g33550	479.87	5.74	1.59	---	lipid transfer protein (LTP) family protein	s		
203	251851_at	At3g54670	477.37	5.75	1.57	---	structural maintenance of chromosomes family	u		
220	245744_at	At1g51110	496.82	5.83	1.7	---	plastid-lipid associated protein PAP family	c		
219	267000_at	At2g34310	496.74	5.85	1.71	---	expressed protein	u		
218	248228_at	At5g53800	496.04	5.85	1.6	---	expressed protein	o		
208	247163_at	At5g65685	485.03	5.85	1.73	---	soluble glycogen synthase-related	c		
207	257822_at	At3g25230	484.19	5.86	1.72	ROF1	peptidyl-prolyl cis-trans isomerise	o		
214	244979_at	rps11	492.32	5.86	1.69	---	Unknown	c		
215	245018_at	ycf4	492.6	5.86	1.58	---	Unknown	c		
217	267268_at	At2g02570	495.88	5.88	1.66	---	expressed protein	u		
206	265627_at	At2g27285	484.15	5.88	1.68	---	expressed protein	u		
213	259511_at	At1g12520	491.87	5.89	1.54	---	superoxide dismutase copper chaperone, putative	c	1	
211	266957_at	At2g34640	490.12	5.89	1.71	---	expressed protein	u		
212	251641_at	At3g57470	491.57	5.89	1.68	---	peptidase M16 / insulinase family protein	o		
216	248032_at	At5g55860	494.74	5.89	1.57	---	expressed protein	o		
221	256880_at	At3g26450	499.57	5.9	1.66	---	major latex protein-related / MLP-related	u		
209	246954_at	At5g04830	488.09	5.9	1.7	---	expressed protein	o		
210	249144_at	At5g43270	490.08	5.92	1.65	SPL2	squamosa promoter-binding protein-like 2 (SPL2)	u		
222	251928_at	At3g53980	502.27	5.96	1.57	---	seed storage/lipid transfer (LTP) family protein	s		
223	260541_at	At2g43530	504.92	6.05	1.76	---	trypsin inhibitor, putative	s		
224	255028_at	At4g09890	508.11	6.14	1.61	---	expressed protein	u		
226	262526_at	At1g17050	511.1	6.18	1.62	---	geranyl diphosphate synthase, putative	u		
225	262299_at	At1g27520	510.9	6.21	1.55	---	glycoside hydrolase family 47 protein	s	4	
227	265902_at	At2g25590	521.08	6.46	1.51	---	agenet domain-containing protein	u		
235	252170_at	At3g50480	528.4	6.48	1.72	---	broad-spectrum mildew resistance family	s		
228	261701_at	At1g32750	523.68	6.5	1.66	HAC13	HAC13 protein (HAC13)	o		
234	261854_at	At1g50670	528.39	6.5	1.68	---	OTU-like cysteine protease family protein	o		
236	262381_at	At1g72900	531.59	6.53	1.54	---	disease resistance protein (TIR-NBS class)	u		
229	251247_at	At3g62140	525.24	6.53	1.62	---	expressed protein	o		
232	255544_at	At4g01880	527.21	6.53	1.55	---	expressed protein	u		
233	253411_at	At4g32980	527.96	6.53	1.69	---	homeobox protein (ATH1)	u		
230	263763_at	At2g21385	526.3	6.55	1.64	---	expressed protein	c		
231	253686_at	At4g29750	526.99	6.55	1.76	---	expressed protein	c		
242	253161_at	At4g35770	543.04	6.78	1.34	SEN1	senescence-associated protein (SEN1)	c		
243	258453_at	At3g22320	544.83	6.79	1.8	---	DNA-directed RNA polymerase, putative	u		
238	258057_at	At3g28970	540.36	6.79	1.55	---	expressed protein	u		
237	245460_at	At4g16990	539.42	6.79	1.7	---	disease resistance protein (TIR-NBS class)	u		

Table B-6 continued. Genes up-regulated in *35S::AtPQL6-1* relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
181	247273_at	At5g64340	439.49	5.02	1.65	---	expressed protein	u		
182	261151_at	At1g19650	443.12	5.13	1.7	---	phosphoglyceride transfer protein, putative	u		
183	266572_at	At2g23840	444.7	5.15	1.64	---	HNH endonuclease domain-containing	u		
184	255088_at	At4g09350	448.53	5.26	1.54	---	DNAJ heat shock protein	u		
188	249308_at	At5g41350	452.98	5.26	1.7	---	zinc finger (C3HC4-type RING finger) family	u		
186	246476_at	At5g16730	451.22	5.27	1.66	---	expressed protein	c		
190	262809_at	At1g11720	455.99	5.28	1.64	---	starch synthase, putative	c		
187	264126_at	At1g79280	452.38	5.28	1.63	---	expressed protein	o		
185	257178_at	At3g13070	450.6	5.3	1.7	---	CBS domain-containing	c		
189	267477_at	At2g02710	455.61	5.31	1.55	---	PAC motif-containing	o		
191	249386_at	At5g40060	457.23	5.31	1.67	---	disease resistance protein (TIR-NBS class)	u		
192	254305_at	At4g22200	459.82	5.38	1.49	<i>AKT2/3</i>	potassium channel protein 2 (AKT2/AKT3)	u		
193	259219_at	At3g03560	461.13	5.39	1.65	---	expressed protein	o		
195	264698_at	At1g70200	463.25	5.44	1.86	---	RNA recognition motif (RRM)-containing	c		
194	254083_at	At4g24920	462.79	5.45	1.89	---	transport protein SEC61 gamma subunit, putative	o		
196	251932_at	At3g54010	464.62	5.47	1.66	<i>PAS1-D</i>	peptidyl-prolyl cis-trans isomerise	o		
198	251182_at	At3g62600	471.23	5.64	1.69	---	DNAJ heat shock family protein	s	1	
197	261886_s_at	At1g80700	471.02	5.66	1.73	---	expressed protein	m		
202	261206_at	At1g12800	475.71	5.68	1.77	---	S1 RNA-binding domain-containing protein	c		
201	261133_at	At1g19720	474.62	5.68	1.55	---	jacalin lectin family	u		
200	265698_at	At2g32160	473.85	5.68	1.67	---	expressed protein	u		
199	249652_at	At5g37070	473.59	5.71	1.64	---	expressed protein	u		
204	247854_at	At5g58200	477.49	5.72	1.65	---	expressed protein	o		
205	253344_at	At4g33550	479.87	5.74	1.59	---	lipid transfer protein (LTP) family protein	s		
203	251851_at	At3g54670	477.37	5.75	1.57	---	structural maintenance of chromosomes family	u		
220	245744_at	At1g51110	496.82	5.83	1.7	---	plastid-lipid associated protein PAP family	c		
219	267000_at	At2g34310	496.74	5.85	1.71	---	expressed protein	u		
218	248228_at	At5g53800	496.04	5.85	1.6	---	expressed protein	o		
208	247163_at	At5g65685	485.03	5.85	1.73	---	soluble glycogen synthase-related	c		
207	257822_at	At3g25230	484.19	5.86	1.72	<i>ROF1</i>	peptidyl-prolyl cis-trans isomerise	o		
214	244979_at	rps11	492.32	5.86	1.69	---	Unknown	c		
215	245018_at	ycf4	492.6	5.86	1.58	---	Unknown	c		
217	267268_at	At2g02570	495.88	5.88	1.66	---	expressed protein	u		
206	265627_at	At2g27285	484.15	5.88	1.68	---	expressed protein	u		
213	259511_at	At1g12520	491.87	5.89	1.54	---	superoxide dismutase copper chaperone, putative	c	1	
211	266957_at	At2g34640	490.12	5.89	1.71	---	expressed protein	u		
212	251641_at	At3g57470	491.57	5.89	1.68	---	peptidase M16 / insulinase family protein	o		
216	248032_at	At5g55860	494.74	5.89	1.57	---	expressed protein	o		
221	256880_at	At3g26450	499.57	5.9	1.66	---	major latex protein-related / MLP-related	u		
209	246954_at	At5g04830	488.09	5.9	1.7	---	expressed protein	o		
210	249144_at	At5g43270	490.08	5.92	1.65	<i>SPL2</i>	squamosa promoter-binding protein-like 2 (SPL2)	u		
222	251928_at	At3g53980	502.27	5.96	1.57	---	seed storage/lipid transfer (LTP) family protein	s		
223	260541_at	At2g43530	504.92	6.05	1.76	---	trypsin inhibitor, putative	s		
224	255028_at	At4g09890	508.11	6.14	1.61	---	expressed protein	u		
226	262526_at	At1g17050	511.1	6.18	1.62	---	geranyl diphosphate synthase, putative	u		
225	262299_at	At1g27520	510.9	6.21	1.55	---	glycoside hydrolase family 47 protein	s	4	
227	265902_at	At2g25590	521.08	6.46	1.51	---	agenet domain-containing protein	u		
235	252170_at	At3g50480	528.4	6.48	1.72	---	broad-spectrum mildew resistance family	s		
228	261701_at	At1g32750	523.68	6.5	1.66	<i>HAC13</i>	HAC13 protein (HAC13)	o		
234	261854_at	At1g50670	528.39	6.5	1.68	---	OTU-like cysteine protease family protein	o		
236	262381_at	At1g72900	531.59	6.53	1.54	---	disease resistance protein (TIR-NBS class)	u		
229	251247_at	At3g62140	525.24	6.53	1.62	---	expressed protein	o		
232	255544_at	At4g01880	527.21	6.53	1.55	---	expressed protein	u		
233	253411_at	At4g32980	527.96	6.53	1.69	---	homeobox protein (ATH1)	u		
230	263763_at	At2g21385	526.3	6.55	1.64	---	expressed protein	c		
231	253686_at	At4g29750	526.99	6.55	1.76	---	expressed protein	c		
242	253161_at	At4g35770	543.04	6.78	1.34	<i>SEN1</i>	senescence-associated protein (SEN1)	c		
243	258453_at	At3g22320	544.83	6.79	1.8	---	DNA-directed RNA polymerase, putative	u		
238	258057_at	At3g28970	540.36	6.79	1.55	---	expressed protein	u		
237	245460_at	At4g16990	539.42	6.79	1.7	---	disease resistance protein (TIR-NBS class)	u		

Table B-6 continued. Genes up-regulated in *35S::AtPQL6-1* relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glycosylation sites	GPI anchoring
241	259765_at	At1g64370	543	6.81	1.77	---	expressed protein	o		
239	248378_at	At5g51840	541.93	6.82	1.61	---	expressed protein	o		
240	245025_at	atpF	542.4	6.82	1.55	---	Unknown	c		
246	247565_at	At5g61150	548.8	6.84	1.57	---	leo1-like family protein	u		
251	255933_at	At1g12750	554.38	6.87	1.5	---	rhomoid family protein	u		
245	247544_at	At5g61670	548.78	6.87	1.62	---	expressed protein	u		
248	246759_at	At5g27950	551.43	6.88	1.62	---	kinesin motor protein-related	u		
247	259533_at	At1g12530	551.15	6.89	1.63	---	hypothetical protein	o		
244	260919_at	At1g21520	548.51	6.89	1.56	---	expressed protein	u		
250	257609_at	At3g13845	552.99	6.89	1.69	---	expressed protein	s		
249	257057_at	At3g15310	552.34	6.9	1.47	---	expressed protein	o		
254	263570_at	At2g27150	559.46	6.94	1.57	AO3	aldehyde oxidase 3 (AAO3)	o		
253	255554_at	At4g01895	558.66	6.95	1.84	---	expressed protein	o		
255	261901_at	At1g80920	560.23	6.96	1.58	---	DNAJ heat shock protein	c		
252	252478_at	At3g46540	557.89	6.96	1.68	---	ENTH domain family protein	u		
261	254052_at	At4g25280	565.42	7.01	1.59	---	adenylate kinase family protein	u		
260	256245_at	At3g12580	564.95	7.02	1.24	---	heat shock protein 70, putative	o		
259	249752_at	At5g24660	563.98	7.03	1.77	---	expressed protein	u		
258	263674_at	At2g04790	563.72	7.06	1.48	---	expressed protein	u		
257	260999_at	At1g26580	563.26	7.08	1.58	---	expressed protein	o		
256	252086_at	At3g52030	562.89	7.08	1.58	---	F-box family protein / WD-40 repeat family protein	o		
262	244938_at	rps15	570.89	7.16	1.64	---	Unknown	c		
264	253422_at	At4g32240	582.64	7.53	1.63	---	expressed protein	u		
265	260048_at	At1g73750	583.92	7.55	1.48	---	expressed protein	u		
266	247062_at	At5g66810	585.02	7.55	1.57	---	expressed protein	u		
263	246275_at	At4g36540	582.5	7.56	1.57	---	basic helix-loop-helix (bHLH) family protein	u		
267	263777_at	At2g46450	586.99	7.6	1.59	CNGC12	cyclic nucleotide-regulated channel	o		
274	262346_at	At1g63980	594.64	7.67	1.74	---	D111/G-patch domain-containing protein	u		
268	265837_at	At2g14560	589.89	7.67	1.66	---	expressed protein	u		
273	261953_at	At1g64440	594.47	7.7	1.58	---	UDP-glucose/galactose 4-epimerase, putative	s		
275	260568_at	At2g43570	596	7.71	1.61	---	chitinase, putative	s	1	
269	258585_at	At3g04340	591.62	7.71	1.63	---	FtsH protease family protein	u		
270	255639_at	At4g00760	592.17	7.72	1.46	---	two-component responsive regulator family protein	u		
272	246214_at	At4g36990	594.38	7.72	1.64	HSF4	heat shock factor protein 4 (HSF4)	u		
271	264633_at	At1g65660	593.49	7.74	1.59	---	zinc knuckle (CCHC-type) family protein	u		
278	252362_at	At3g48500	599.88	7.78	1.68	---	expressed protein	c		
277	248807_at	At5g47500	598.99	7.78	1.65	---	pectinesterase family protein	s	4	
276	253048_at	At4g37560	598.25	7.79	1.52	---	formamidase, putative	u		
279	251720_at	At3g56160	601.9	7.86	1.47	---	expressed protein	c		
280	254659_at	At4g18240	603.47	7.86	1.67	---	starch synthase-related protein	c		
281	258133_at	At3g24500	604.67	7.87	1.51	---	ethylene-responsive transcriptional coactivator	u		
282	263972_at	At2g42760	606.67	7.9	1.53	---	expressed protein	o		
289	255836_at	At2g33440	613.09	7.92	1.51	---	splicing factor family protein	u		
288	257285_at	At3g29760	612.5	7.92	1.56	---	NLI interacting factor (NIF) family protein	o		
287	248743_at	At5g48240	612.49	7.95	1.72	---	hypothetical protein	u		
286	248471_at	At5g50840	611.94	7.97	1.73	---	expressed protein	u		
292	263039_at	At1g23280	617.14	8	1.64	---	MAK16 protein-related	u		
283	256805_at	At3g20930	610.5	8	1.68	---	RNA recognition motif (RRM)-containing protein	c		
285	249754_at	At5g24530	611.91	8	1.5	---	2OG-Fe(II) oxygenase family	u		
290	253597_at	At4g30690	616.24	8.01	1.65	IF-3	translation initiation factor 3 family	c		
284	248694_at	At5g48340	611.27	8.01	1.55	---	expressed protein	u		
291	262161_at	At1g52600	616.9	8.02	1.63	---	signal peptidase, putative	u		
298	260235_at	At1g74560	622.42	8.02	1.58	---	nucleosome assembly protein family	o		
297	258880_at	At3g06420	622.08	8.02	1.53	APG8h	autophagy 8h protein	o		
293	261891_at	At1g80790	618.34	8.03	1.63	---	XH/XS domain-containing protein	o		
296	258456_at	At3g22420	621.52	8.04	1.46	---	protein kinase family protein	o		
295	249320_at	At5g40910	621.13	8.05	1.48	---	disease resistance protein (TIR-NBS class)	u		
294	264572_at	At1g05320	620.45	8.07	1.62	---	myosin-related	o		
299	266778_at	At2g29090	626.41	8.1	1.49	---	cytochrome P450 family protein	s		
300	246840_at	At5g26605	630.55	8.19	1.61	---	D111/G-patch domain-containing protein	o		

Table B-6 continued. Genes up-regulated in *35S::AtPQL6-1* relative to wild type.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glyc. sites	GPI anchor
302	249711_at	At5g35680	634.41	8.28	1.59	<i>ELF-1A</i>	translation initiation factor 1A, putative	o		
301	248913_at	At5g45760	633.88	8.29	1.63	---	WD-40 repeat family protein	u		
305	247738_at	At5g59210	640.11	8.38	1.59	---	myosin heavy chain-related	u		
303	267518_at	At2g30500	638.57	8.39	1.55	---	kinase interacting family protein	u		
304	251790_at	At3g55470	639.55	8.4	1.63	---	C2 domain-containing protein	u		
306	246311_at	At3g51880	641.82	8.41	1.61	<i>HMGa</i>	high mobility group protein alpha	o		
307	247521_at	At5g61450	644.17	8.45	1.52	---	2-phosphoglycerate kinase-related	c		
310	253068_at	At4g37820	647.08	8.46	1.53	---	expressed protein	u		
308	263802_at	At2g40430	645.55	8.47	1.68	---	expressed protein	u		
309	245017_at	psal	646.83	8.48	1.66	---	Unknown	c		
311	245219_at	At1g59124	650.66	8.55	1.49	---	disease resistance protein (CC-NBS-LRR class)	u		
313	267483_at	At2g02810	652.59	8.57	1.6	<i>UTR1</i>	UDP-galactose/UDP-glucose transporter	s		
312	246997_at	At5g67390	651.8	8.57	1.9	---	expressed protein	o		
314	245533_at	At4g15130	654.09	8.58	1.67	---	phosphorylcholine transferase, putative	o		
316	250786_at	At5g05540	656.13	8.62	1.54	---	exonuclease family protein	u		
315	256676_at	At3g52180	656.13	8.65	1.63	<i>PTPKIS1</i>	protein tyrosine phosphatase/kinase interaction	u		
317	267257_at	At2g23080	659.94	8.75	1.49	---	casein kinase II alpha chain, putative	u		
320	262288_at	At1g70760	662.91	8.78	1.56	---	inorganic carbon transport protein-related	u		
319	246069_at	At5g20220	662.42	8.79	1.52	---	zinc knuckle (CCHC-type) family protein	u		
318	263584_at	At2g17040	661.56	8.8	1.62	---	no apical meristem (NAM) family protein	o		
321	255913_at	At1g66980	665.7	8.87	1.47	---	glycerophosphodiesterase family	u	14	
322	259417_at	At1g02340	667.2	8.9	1.44	<i>FBI1</i>	basic helix-loop-helix protein (FBI1)	u		
323	266727_at	At2g03150	667.58	8.9	1.57	---	ATP/GTP-binding protein family	m		
324	267089_at	At2g38300	668.32	8.9	1.51	---	myb family transcription factor	o		
326	254596_at	At4g18975	673.07	9.02	1.54	---	pentatricopeptide (PPR) repeat-containing protein	u		
325	248031_at	At5g55640	672.35	9.02	1.57	---	expressed protein	o		
327	259520_at	At1g12320	676.12	9.13	1.44	---	expressed protein	o		
328	254911_at	At4g11100	676.9	9.13	1.62	---	expressed protein	o		
329	248821_at	At5g47070	681.46	9.26	1.63	---	protein kinase, putative	u		
332	257628_at	At3g26290	686	9.3	1.57	---	cytochrome P450 71B26, putative (CYP71B26)	s	2	
331	252025_at	At3g52900	684.62	9.31	1.54	---	expressed protein	u		
333	249125_at	At5g43450	687.16	9.31	1.55	---	2-oxoglutarate-dependent dioxygenase, putative	u		
330	248527_at	At5g50740	683.61	9.31	1.66	---	copper chaperone (CCH)-related	o		
335	261644_s_at	At1g27830	689.48	9.34	1.51	---	WD-40 repeat family protein	o		
338	265025_at	At1g24575	691.76	9.36	1.61	---	expressed protein	u		
334	255566_s_at	At4g01780	689.31	9.36	1.62	---	XH/XS domain-containing protein	u		
336	253687_at	At4g29520	690.82	9.37	1.52	---	expressed protein	s		
337	258291_at	At3g23310	691.75	9.39	1.55	---	protein kinase, putative	u		
339	246818_at	At5g27270	694.85	9.45	1.51	---	pentatricopeptide (PPR) repeat-containing protein	c		
341	260336_at	At1g73990	696.93	9.46	1.6	<i>SppA</i>	peptidase U7 family protein	c		
340	254661_at	At4g18260	696.73	9.48	1.51	---	cytochrome B561-related	s	2	
344	256980_at	At3g26932	701.78	9.54	1.54	---	DsRBD domain-containing protein	u		
343	249811_at	At5g23760	701.06	9.55	1.56	---	heavy-metal-associated domain-containing protein	u		
345	261800_at	At1g30490	703.57	9.57	1.5	<i>Athb-9</i>	leucine zipper transcription factor	o		
342	247820_at	At5g58380	700.94	9.57	1.61	---	CBL-interacting protein kinase 10 (CIPK10)	u		
350	247340_at	At5g63700	711.94	9.69	1.49	---	zinc finger (C3HC4 type RING finger) family protein	o		
349	257665_at	At3g20430	711.78	9.7	1.62	---	expressed protein	o		
348	250310_at	At5g12230	710.6	9.7	1.54	---	expressed protein	u		
351	249904_at	At5g22700	713.34	9.7	1.52	---	F-box family protein	u		
354	256001_at	At1g29850	716.92	9.72	1.61	---	double-stranded DNA-binding family protein	u		
347	259295_at	At3g05340	710.37	9.72	1.6	---	pentatricopeptide (PPR) repeat-containing protein	u		
352	249299_at	At5g41340	715.59	9.72	1.51	<i>UBC4</i>	ubiquitin-conjugating enzyme 4 (UBC4)	o		
353	264580_at	At1g05340	716.51	9.73	1.42	---	expressed protein	c		
346	266971_at	At2g39580	709.83	9.73	1.52	---	expressed protein	u		
355	250571_at	At5g08200	719.39	9.78	1.56	---	peptidoglycan-binding LysM domain protein	u		
357	246479_at	At5g16060	722.57	9.84	1.5	---	expressed protein	u		
358	250064_at	At5g17890	723.39	9.85	1.52	---	LIM domain-containing protein	u		
356	245049_at	rps16	721.87	9.85	1.66	---	Unknown	c		
360	263222_at	At1g30640	725.24	9.86	1.57	---	protein kinase, putative	u		
359	262693_at	At1g62780	724.12	9.86	1.8	---	expressed protein	c		

Table B-7. Genes down-regulated in *35S::AtPQL6-I* relative to wild type. Annotation as for B-1.

Rank	Probe	AGI Accession	RP Score	FDR (%)	Fold Change	Gene Name	Description	Localisation	N-glycosylation sites	GPI anchoring
1	248844_s_at	At5g46900	14.55	0	-2.59	---	seed storage/lipid transfer protein (LTP) family protein	s		
2	265117_at	At1g62500	49.68	0.5	-1.83	---	seed storage/lipid transfer protein (LTP) family protein	s		
3	267411_at	At2g34930	92.28	3.67	-1.76	---	disease resistance family protein	s	7	
4	260869_at	At1g43800	116.62	6	-1.74	---	acyl-(acyl-carrier-protein) desaturase, putative	u		
5	254820_s_at	At4g12510	125.02	6	-1.7	---	seed storage/lipid transfer protein (LTP) family protein	s		
7	260914_at	At1g02640	146.38	6.14	-1.65	---	glycosyl hydrolase family 3 protein	s	1	
10	246002_at	At5g20740	171.02	6.3	-1.55	---	invertase/pectin methylesterase inhibitor family protein	s		
6	244959_s_at	orf107c	139.68	6.33	-1.64	---	expressed protein	m		
9	264824_at	At1g03420	168.98	6.89	-1.54	---	expressed protein	o		
8	248614_at	At5g49560	157.97	7	-1.54	---	expressed protein	o		
13	261658_at	At1g50040	195.93	7.15	-1.56	---	expressed protein	c		
15	245325_at	At4g14130	207.37	7.27	-1.58	<i>XTR7</i>	endo-xyloglucan transferase, putative (XTR7)	s	1	
12	254770_at	At4g13340	193.43	7.33	-1.61	---	leucine-rich repeat / extensin family protein	s	5	
17	263443_at	At2g28630	218.95	7.53	-1.5	---	beta-ketoacyl-CoA synthase family protein	s	3	
16	244973_at	psbT	216.86	7.69	-1.43	---	Unknown	c		
14	264953_at	At1g77120	207.03	7.79	-1.54	<i>ADH</i>	alcohol dehydrogenase (ADH)	u		
18	248727_at	At5g47990	227.92	7.83	-1.51	---	cytochrome P450 family protein	s	4	
11	266460_at	At2g47930	191.96	8	-1.59	---	hydroxyproline-rich glycoprotein family protein	u		Yes
21	258005_at	At3g19390	256.63	9.1	-1.47	---	cysteine proteinase, putative / thiol protease, putative	s	2	
20	245928_s_at	At5g24780	254.25	9.2	-1.56	<i>VSP2</i>	vegetative storage protein 2 (VSP2)	s	1	
19	264857_at	At1g24170	252.44	9.58	-1.58	---	glycosyl transferase family 8 protein	s	1	

References

- Alabadi, D., Oyama, T., Yanovsky, M.J., Harmon, F.G., Mas, P., and Kay, S.A.** (2001). Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science* **293**, 880-883.
- Alberts, B.** (2002). *Molecular biology of the cell*. (New York: Garland Science).
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J.** (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.
- Alvim, F.C., Carolino, S.M., Cascardo, J.C., Nunes, C.C., Martinez, C.A., Otoni, W.C., and Fontes, E.P.** (2001). Enhanced accumulation of BiP in transgenic plants confers tolerance to water stress. *Plant Physiol* **126**, 1042-1054.
- Anand, M., Rush, J.S., Ray, S., Doucey, M.A., Weik, J., Ware, F.E., Hofsteenge, J., Waechter, C.J., and Lehrman, M.A.** (2001). Requirement of the Lec35 gene for all known classes of monosaccharide-P-dolichol-dependent glycosyltransferase reactions in mammals. *Mol Biol Cell* **12**, 487-501.
- Anantharaman, V., and Aravind, L.** (2003). Application of comparative genomics in the identification and analysis of novel families of membrane-associated receptors in bacteria. *BMC Genomics* **4**, 34.
- Apone, F., Alyeshmerni, N., Wiens, K., Chalmers, D., Chrispeels, M.J., and Colucci, G.** (2003). The G-protein-coupled receptor GCR1 regulates DNA synthesis through activation of phosphatidylinositol-specific phospholipase C. *Plant Physiol* **133**, 571-579.
- Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M.D., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N.J., Oinn, T.M., Pagni, M., Servant, F., Sigrist, C.J., and Zdobnov, E.M.** (2000). InterPro--an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* **16**, 1145-1150.
- Arteca, R.N., and Arteca, J.M.** (2000). A novel method for growing Arabidopsis thaliana plants hydroponically. *Physiol Plantarum* **108**, 188-193.
- Attard, M., Jean, G., Forestier, L., Cherqui, S., van't Hoff, W., Broyer, M., Antignac, C., and Town, M.** (1999). Severity of phenotype in cystinosis varies with mutations in the CTNS gene: predicted effect on the model of cystinosis. *Hum Mol Genet* **8**, 2507-2514.
- Aviezer-Hagai, K., Skovorodnikova, J., Galigniana, M., Farchi-Pisanty, O., Maayan, E., Bocovza, S., Efrat, Y., von Koskull-Doring, P., Ohad, N., and Breiman, A.** (2007). Arabidopsis immunophilins ROF1 (AtFKBP62) and ROF2 (AtFKBP65) exhibit tissue specificity, are heat-stress induced, and bind HSP90. *Plant Mol Biol* **63**, 237-255.
- Baldwin, T.C., Handford, M.G., Yuseff, M.I., Orellana, A., and Dupree, P.** (2001). Identification and characterization of GONST1, a golgi-localized GDP-mannose transporter in Arabidopsis. *Plant Cell* **13**, 2283-2295.
- Betts, M.R., and Russell, R.B.** (2003). Amino Acid Properties and Consequences of Substitutions. in *Bioinformatics for Geneticists* M.R. Barnes, I.C. Gray (eds).

- Boisson, M., Gomord, V., Audran, C., Berger, N., Dubreucq, B., Granier, F., Lerouge, P., Faye, L., Caboche, M., and Lepiniec, L.** (2001). Arabidopsis glucosidase I mutants reveal a critical role of N-glycan trimming in seed development. *Embo J* **20**, 1010-1019.
- Borner, G.H., Lilley, K.S., Stevens, T.J., and Dupree, P.** (2003). Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. *Plant Physiol* **132**, 568-577.
- Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., and Grolach, J.** (2001). Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell* **13**, 1499-1510.
- Brady, S.M., Song, S., Dhugga, K.S., Rafalski, J.A., and Benfey, P.N.** (2007). Combining expression and comparative evolutionary analysis. The COBRA gene family. *Plant Physiol* **143**, 172-187.
- Brazier-Hicks, M., Offen, W.A., Gershtater, M.C., Revett, T.J., Lim, E.K., Bowles, D.J., Davies, G.J., and Edwards, R.** (2007). Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc Natl Acad Sci U S A* **104**, 20238-20243.
- Breitling, R., Amtmann, A., and Herzyk, P.** (2004a). Iterative Group Analysis (iGA): a simple tool to enhance sensitivity and facilitate interpretation of microarray experiments. *BMC Bioinformatics* **5**, 34.
- Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P.** (2004b). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* **573**, 83-92.
- Brett, C.L., Tukaye, D.N., Mukherjee, S., and Rao, R.** (2005). The yeast endosomal Na⁺/K⁺/H⁺ exchanger Nhx1 regulates cellular pH to control vesicle trafficking. *Mol Biol Cell* **16**, 1396-1405.
- Brett, C.T., and Leloir, L.F.** (1977). Dolichyl monophosphate and its sugar derivatives in plants. *Biochem J* **161**, 93-101.
- Burda, P., te Heesen, S., Brachat, A., Wach, A., Dusterhoft, A., and Aebi, M.** (1996). Stepwise assembly of the lipid-linked oligosaccharide in the endoplasmic reticulum of *Saccharomyces cerevisiae*: identification of the ALG9 gene encoding a putative mannosyl transferase. *Proc Natl Acad Sci U S A* **93**, 7160-7165.
- Burn, J.E., Hurley, U.A., Birch, R.J., Arioli, T., Cork, A., and Williamson, R.E.** (2002). The cellulose-deficient Arabidopsis mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. *Plant J* **32**, 949-960.
- Callis, J.** (1995). Regulation of Protein Degradation. *Plant Cell* **7**, 845-857.
- Cameron, K.D., Teece, M.A., and Smart, L.B.** (2006). Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. *Plant Physiol* **140**, 176-183.
- Camp, L.A., Chauhan, P., Farrar, J.D., and Lehrman, M.A.** (1993). Defective mannosylation of glycosylphosphatidylinositol in Lec35 Chinese hamster ovary cells. *J Biol Chem* **268**, 6721-6728.

- Campbell, P., and Braam, J.** (1999). Xyloglucan endotransglycosylases: diversity of genes, enzymes and potential wall-modifying functions. *Trends Plant Sci* **4**, 361-366.
- Chadwick, C.M., and Northcote, D.H.** (1980). Glucosylation of phosphorylpolyisoprenol and sterol at the plasma membrane of soya-bean (*Glycine max*) protoplasts. *Biochem J* **186**, 411-421.
- Chang, E.C., Barr, M., Wang, Y., Jung, V., Xu, H.P., and Wigler, M.H.** (1994). Cooperative interaction of *S. pombe* proteins required for mating and morphogenesis. *Cell* **79**, 131-141.
- Chen, J.G., Willard, F.S., Huang, J., Liang, J., Chasse, S.A., Jones, A.M., and Siderovski, D.P.** (2003). A seven-transmembrane RGS protein that modulates plant cell proliferation. *Science* **301**, 1728-1731.
- Chen, J.G., Pandey, S., Huang, J., Alonso, J.M., Ecker, J.R., Assmann, S.M., and Jones, A.M.** (2004). GCR1 can act independently of heterotrimeric G-protein in response to brassinosteroids and gibberellins in *Arabidopsis* seed germination. *Plant Physiol* **135**, 907-915.
- Chen, S., Sanchez-Fernandez, R., Lyver, E.R., Dancis, A., and Rea, P.A.** (2007). Functional characterization of AtATM1, AtATM2, and AtATM3, a subfamily of *Arabidopsis* half-molecule ATP-binding cassette transporters implicated in iron homeostasis. *J Biol Chem* **282**, 21561-21571.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D.** (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* **31**, 3497-3500.
- Cherqui, S., Kalatzis, V., Trugnan, G., and Antignac, C.** (2001). The targeting of cystinosis to the lysosomal membrane requires a tyrosine-based signal and a novel sorting motif. *J Biol Chem* **276**, 13314-13321.
- Cherqui, S., Kalatzis, V., Forestier, L., Poras, I., and Antignac, C.** (2000). Identification and characterisation of the murine homologue of the gene responsible for cystinosis, Ctns. *BMC Genomics* **1**, 2.
- Chol, M., Nevo, N., Cherqui, S., Antignac, C., and Rustin, P.** (2004). Glutathione precursors replenish decreased glutathione pool in cystinotic cell lines. *Biochem Biophys Res Commun* **324**, 231-235.
- Chung, K.S., Hoe, K.L., Kim, K.W., and Yoo, H.S.** (1998). Isolation of a novel heat shock protein 70-like gene, pss1+ of *Schizosaccharomyces pombe* homologous to hsp110/SSE subfamily. *Gene* **210**, 143-150.
- Chung, K.S., Won, M., Lee, S.B., Jang, Y.J., Hoe, K.L., Kim, D.U., Lee, J.W., Kim, K.W., and Yoo, H.S.** (2001). Isolation of a novel gene from *Schizosaccharomyces pombe*: stm1+ encoding a seven-transmembrane loop protein that may couple with the heterotrimeric Galpha 2 protein, Gpa2. *J Biol Chem* **276**, 40190-40201.
- Clamp, M., Cuff, J., Searle, S.M., and Barton, G.J.** (2004). The Jalview Java alignment editor. *Bioinformatics* **20**, 426-427.
- Clough, S.J., and Bent, A.F.** (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**, 735-743.
- Cobbett, C.S.** (2000). Phytochelatin and their roles in heavy metal detoxification. *Plant Physiol* **123**, 825-832.

- Colucci, G., Apone, F., Alyeshmerni, N., Chalmers, D., and Chrispeels, M.J.** (2002). GCR1, the putative Arabidopsis G protein-coupled receptor gene is cell cycle-regulated, and its overexpression abolishes seed dormancy and shortens time to flowering. *Proc Natl Acad Sci U S A* **99**, 4736-4741.
- Conboy, M.J., and Cyert, M.S.** (2000). Luv1p/Rki1p/Tcs3p/Vps54p, a yeast protein that localizes to the late Golgi and early endosome, is required for normal vacuolar morphology. *Mol Biol Cell* **11**, 2429-2443.
- Conklin, P.L., Williams, E.H., and Last, R.L.** (1996). Environmental stress sensitivity of an ascorbic acid-deficient Arabidopsis mutant. *Proc Natl Acad Sci U S A* **93**, 9970-9974.
- Conklin, P.L., Norris, S.R., Wheeler, G.L., Williams, E.H., Smirnoff, N., and Last, R.L.** (1999). Genetic evidence for the role of GDP-mannose in plant ascorbic acid (vitamin C) biosynthesis. *Proc Natl Acad Sci U S A* **96**, 4198-4203.
- Cooper, A.J.** (1983). Biochemistry of sulfur-containing amino acids. *Annu Rev Biochem* **52**, 187-222.
- Cox, J.S., Chapman, R.E., and Walter, P.** (1997). The unfolded protein response coordinates the production of endoplasmic reticulum protein and endoplasmic reticulum membrane. *Mol Biol Cell* **8**, 1805-1814.
- Cunillera, N., Arro, M., Fores, O., Manzano, D., and Ferrer, A.** (2000). Characterization of dehydrodolichyl diphosphate synthase of Arabidopsis thaliana, a key enzyme in dolichol biosynthesis. *FEBS Lett* **477**, 170-174.
- Cutler, S.R., Ehrhardt, D.W., Griffiths, J.S., and Somerville, C.R.** (2000). Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. *Proc Natl Acad Sci U S A* **97**, 3718-3723.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R.** (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* **139**, 5-17.
- D'Amico, L., Valsasina, B., Daminati, M.G., Fabbrini, M.S., Nitti, G., Bollini, R., Ceriotti, A., and Vitale, A.** (1992). Bean homologs of the mammalian glucose-regulated proteins: induction by tunicamycin and interaction with newly synthesized seed storage proteins in the endoplasmic reticulum. *Plant J* **2**, 443-455.
- Danon, A., Rotari, V.I., Gordon, A., Mailhac, N., and Gallois, P.** (2004). Ultraviolet-C overexposure induces programmed cell death in Arabidopsis, which is mediated by caspase-like activities and which can be suppressed by caspase inhibitors, p35 and Defender against Apoptotic Death. *Journal of Biological Chemistry* **279**, 779-787.
- De Vries, L., Zheng, B., Fischer, T., Elenko, E., and Farquhar, M.G.** (2000). The regulator of G protein signaling family. *Annu Rev Pharmacol Toxicol* **40**, 235-271.
- Dean, N.** (1995). Yeast glycosylation mutants are sensitive to aminoglycosides. *Proc Natl Acad Sci U S A* **92**, 1287-1291.
- Divol, F., Vilaine, F., Thibivilliers, S., Kusiak, C., Sauge, M.H., and Dinant, S.** (2007). Involvement of the xyloglucan endotransglycosylase/hydrolases encoded by celery XTH1 and Arabidopsis XTH33 in the phloem response to aphids. *Plant Cell Environ* **30**, 187-201.

- Dominguez-Solis, J.R., Gutierrez-Alcala, G., Vega, J.M., Romero, L.C., and Gotor, C.** (2001). The cytosolic O-acetylserine(thiol)lyase gene is regulated by heavy metals and can function in cadmium tolerance. *J Biol Chem* **276**, 9297-9302.
- Doucey, M.A., Hess, D., Cacan, R., and Hofsteenge, J.** (1998). Protein C-mannosylation is enzyme-catalysed and uses dolichyl-phosphate-mannose as a precursor. *Mol Biol Cell* **9**, 291-300.
- Downing, W.L., Galpin, J.D., Clemens, S., Lauzon, S.M., Samuels, A.L., Pidkowich, M.S., Clarke, L.A., and Kermode, A.R.** (2006). Synthesis of enzymatically active human alpha-L-iduronidase in *Arabidopsis* cgl (complex glycan-deficient) seeds. *Plant Biotechnol J* **4**, 169-181.
- Dragon, F., Gallagher, J.E., Compagnone-Post, P.A., Mitchell, B.M., Porwancher, K.A., Wehner, K.A., Wormsley, S., Settlage, R.E., Shabanowitz, J., Osheim, Y., Beyer, A.L., Hunt, D.F., and Baserga, S.J.** (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**, 967-970.
- Edwards, K., Johnstone, C., and Thompson, C.** (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res* **19**, 1349.
- Eisenhaber, B., Wildpaner, M., Schultz, C.J., Borner, G.H., Dupree, P., and Eisenhaber, F.** (2003). Glycosylphosphatidylinositol lipid anchoring of plant proteins. Sensitive prediction from sequence- and genome-wide studies for *Arabidopsis* and rice. *Plant Physiol* **133**, 1691-1701.
- Elbein, A.D.** (1987). Inhibitors of the biosynthesis and processing of N-linked oligosaccharide chains. *Annu Rev Biochem* **56**, 497-534.
- Ellis, C., and Turner, J.G.** (2002). A conditionally fertile *coi1* allele indicates cross-talk between plant hormone signalling pathways in *Arabidopsis thaliana* seeds and young seedlings. *Planta* **215**, 549-556.
- Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G.** (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* **300**, 1005-1016.
- Fan, L.M., Wang, Y.F., Wang, H., and Wu, W.H.** (2001). In vitro *Arabidopsis* pollen germination and characterization of the inward potassium currents in *Arabidopsis* pollen grain protoplasts. *J Exp Bot* **52**, 1603-1614.
- Faye, L., and Chrispeels, M.J.** (1989). Apparent Inhibition of beta-Fructosidase Secretion by Tunicamycin May Be Explained by Breakdown of the Unglycosylated Protein during Secretion. *Plant Physiol* **89**, 845-851.
- Finn, R.D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S.R., Sonnhammer, E.L., and Bateman, A.** (2006). Pfam: clans, web tools and services. *Nucleic Acids Res* **34**, D247-251.
- Fluckiger, R., De Caroli, M., Piro, G., Dalessandro, G., Neuhaus, J.M., and Di Sansebastiano, G.P.** (2003). Vacuolar system distribution in *Arabidopsis* tissues, visualized using GFP fusion proteins. *J Exp Bot* **54**, 1577-1584.
- Frugoli, J.A., McPeck, M.A., Thomas, T.L., and McClung, C.R.** (1998). Intron loss and gain during evolution of the catalase gene family in angiosperms. *Genetics* **149**, 355-365.

- Fukui, Y., Kozasa, T., Kaziro, Y., Takeda, T., and Yamamoto, M.** (1986). Role of a ras homolog in the life cycle of *Schizosaccharomyces pombe*. *Cell* **44**, 329-336.
- Furmanek, A., and Hofsteenge, J.** (2000). Protein C-mannosylation: facts and questions. *Acta Biochim Pol* **47**, 781-789.
- Gahl, W.A., Thoene, J.G., and Schneider, J.A.** (2002). Cystinosis. *N Engl J Med* **347**, 111-121.
- Gao, N., and Lehrman, M.A.** (2002). Analyses of dolichol pyrophosphate-linked oligosaccharides in cell cultures and tissues by fluorophore-assisted carbohydrate electrophoresis. *Glycobiology* **12**, 353-360.
- Gao, N., Shang, J., and Lehrman, M.A.** (2007a). Unexpected Basis For Impaired Glc3Man9GlcNAc2-P-P-Dolichol Biosynthesis by Elevated Expression of Glcnac-1-p Transferase. *Glycobiology*.
- Gao, X.D., Wang, J., Keppler-Ross, S., and Dean, N.** (2005). ERS1 encodes a functional homologue of the human lysosomal cystine transporter. *Febs J* **272**, 2497-2511.
- Gao, Y., Zeng, Q., Guo, J., Cheng, J., Ellis, B.E., and Chen, J.G.** (2007b). Genetic characterization reveals no role for the reported ABA receptor, GCR2, in ABA control of seed germination and early seedling development in *Arabidopsis*. *Plant J.*
- Gietz, R.D., Schiestl, R.H., Willems, A.R., and Woods, R.A.** (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355-360.
- Gillmor, C.S., Poindexter, P., Lorieau, J., Palcic, M.M., and Somerville, C.** (2002). Alpha-glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*. *J Cell Biol* **156**, 1003-1013.
- Gillmor, C.S., Lukowitz, W., Brininstool, G., Sedbrook, J.C., Hamann, T., Poindexter, P., and Somerville, C.** (2005). Glycosylphosphatidylinositol-anchored proteins are required for cell wall synthesis and morphogenesis in *Arabidopsis*. *Plant Cell* **17**, 1128-1140.
- Gonzalez, A., Jimenez, A., Vazquez, D., Davies, J.E., and Schindler, D.** (1978). Studies on the mode of action of hygromycin B, an inhibitor of translocation in eukaryotes. *Biochim Biophys Acta* **521**, 459-469.
- Greenberg, J.T.** (1997). Programmed Cell Death in Plant-Pathogen Interactions. *Annu Rev Plant Physiol Plant Mol Biol* **48**, 525-545.
- Grobler, J., Bauer, F., Subden, R.E., and Van Vuuren, H.J.** (1995). The *mae1* gene of *Schizosaccharomyces pombe* encodes a permease for malate and other C4 dicarboxylic acids. *Yeast* **11**, 1485-1491.
- Hall, S.E., Kettler, G., and Preuss, D.** (2003). Centromere satellites from *Arabidopsis* populations: maintenance of conserved and variable domains. *Genome Res* **13**, 195-205.
- Hamamoto, A., and Mazelis, M.** (1986). The C-S Lyases of Higher Plants : Isolation and Properties of Homogeneous Cystine Lyase from Broccoli (*Brassica oleracea* var *botrytis*) Buds. *Plant Physiol* **80**, 702-706.
- Handford, M.G., Sicilia, F., Brandizzi, F., Chung, J.H., and Dupree, P.** (2004). *Arabidopsis thaliana* expresses multiple Golgi-localised nucleotide-sugar transporters related to GONST1. *Mol Genet Genomics* **272**, 397-410.

- Hara, K., Yagi, M., Koizumi, N., Kusano, T., and Sano, H.** (2000). Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants. *Plant Cell Physiol* **41**, 684-691.
- Hardwick, K.G., and Pelham, H.R.** (1990). ERS1 a seven transmembrane domain protein from *Saccharomyces cerevisiae*. *Nucleic Acids Res* **18**, 2177.
- Hardwick, K.G., Lewis, M.J., Semenza, J., Dean, N., and Pelham, H.R.** (1990). ERD1, a yeast gene required for the retention of luminal endoplasmic reticulum proteins, affects glycoprotein processing in the Golgi apparatus. *Embo J* **9**, 623-630.
- Harigaya, Y., and Yamamoto, M.** (2007). Molecular mechanisms underlying the mitosis-meiosis decision. *Chromosome Res* **15**, 523-537.
- Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A., and Kay, S.A.** (2000). Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* **290**, 2110-2113.
- Hasegawa, P.M., Bressan, R.A., Zhu, J.K., and Bohnert, H.J.** (2000). Plant Cellular and Molecular Responses to High Salinity. *Annu Rev Plant Physiol Plant Mol Biol* **51**, 463-499.
- Haseloff, J., Siemering, K.R., Prasher, D.C., and Hodge, S.** (1997). Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc Natl Acad Sci U S A* **94**, 2122-2127.
- He, D.Y., Yazaki, Y., Nishizawa, Y., Takai, R., Yamada, K., Sakano, K., Shibuya, N., and Minami, E.** (1998). Gene activation by cytoplasmic acidification in suspension-cultured rice cells in response to the potent elicitor, N-acetylchitoheptaose. *Mol Plant Microbe In* **11**, 1167-1174.
- He, Y., Tang, R.H., Hao, Y., Stevens, R.D., Cook, C.W., Ahn, S.M., Jing, L., Yang, Z., Chen, L., Guo, F., Fiorani, F., Jackson, R.B., Crawford, N.M., and Pei, Z.M.** (2004). Nitric oxide represses the *Arabidopsis* floral transition. *Science* **305**, 1968-1971.
- Helenius, A., and Aebi, M.** (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73**, 1019-1049.
- Helenius, J., Ng, D.T., Marolda, C.L., Walter, P., Valvano, M.A., and Aebi, M.** (2002). Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* **415**, 447-450.
- Henikoff, S., and Henikoff, J.G.** (1992). Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* **89**, 10915-10919.
- Heslop-Harrison, J., and Heslop-Harrison, Y.** (1970). Evaluation of pollen viability by enzymatically induced fluorescence; intracellular hydrolysis of fluorescein diacetate. *Stain Technol* **45**, 115-120.
- Hirschi, K.D.** (1999). Expression of *Arabidopsis* CAX1 in tobacco: altered calcium homeostasis and increased stress sensitivity. *Plant Cell* **11**, 2113-2122.
- Hoeberichts, F.A., Vaeck, E., Kiddle, G., Coppens, E., van de Cotte, B., Adamantidis, A., Ormenese, S., Foyer, C.H., Zabeau, M., Inze, D., Perilleux, C., Van Breusegem, F., and Vuylsteke, M.** (2008). A temperature-sensitive

- mutation in the *Arabidopsis thaliana* phosphomannomutase gene disrupts protein glycosylation and triggers cell death. *J Biol Chem* **283**, 5708-5718.
- Hoffman, C.S.** (2005). Except in every detail: comparing and contrasting G-protein signaling in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Eukaryot Cell* **4**, 495-503.
- Hofsteenge, J., Muller, D.R., de Beer, T., Loffler, A., Richter, W.J., and Vliegenthart, J.F.** (1994). New type of linkage between a carbohydrate and a protein: C-glycosylation of a specific tryptophan residue in human RNase Us. *Biochemistry* **33**, 13524-13530.
- Hollenbach, B., Schreiber, L., Hartung, W., and Dietz, K.J.** (1997). Cadmium leads to stimulated expression of the lipid transfer protein genes in barley: implications for the involvement of lipid transfer proteins in wax assembly. *Planta* **203**, 9-19.
- Honys, D., and Twell, D.** (2004). Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol* **5**, R85.
- Huang, J., Taylor, J.P., Chen, J.G., Uhrig, J.F., Schnell, D.J., Nakagawa, T., Korth, K.L., and Jones, A.M.** (2006). The plastid protein THYLAKOID FORMATION1 and the plasma membrane G-protein GPA1 interact in a novel sugar-signaling mechanism in *Arabidopsis*. *Plant Cell* **18**, 1226-1238.
- Huffaker, T.C., and Robbins, P.W.** (1982). Temperature-sensitive yeast mutants deficient in asparagine-linked glycosylation. *J Biol Chem* **257**, 3203-3210.
- Ichikawa, T., Nakazawa, M., Kawashima, M., Muto, S., Gohda, K., Suzuki, K., Ishikawa, A., Kobayashi, H., Yoshizumi, T., Tsumoto, Y., Tshara, Y., Iizumi, H., Goto, Y., and Matsui, M.** (2003). Sequence database of 1172 T-DNA insertion sites in *Arabidopsis* activation-tagging lines that showed phenotypes in T1 generation. *Plant J* **36**, 421-429.
- Ilgoutz, S.C., Zawadzki, J.L., Ralton, J.E., and McConville, M.J.** (1999). Evidence that free GPI glycolipids are essential for growth of *Leishmania mexicana*. *Embo J* **18**, 2746-2755.
- Innocenti, G., Pucciariello, C., Le Gleuher, M., Hopkins, J., de Stefano, M., Delledonne, M., Puppo, A., Baudouin, E., and Frendo, P.** (2007). Glutathione synthesis is regulated by nitric oxide in *Medicago truncatula* roots. *Planta* **225**, 1597-1602.
- Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P.** (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264.
- Isshiki, T., Mochizuki, N., Maeda, T., and Yamamoto, M.** (1992). Characterization of a fission yeast gene, *gpa2*, that encodes a G alpha subunit involved in the monitoring of nutrition. *Genes Dev* **6**, 2455-2462.
- Iwata, Y., and Koizumi, N.** (2005a). Unfolded protein response followed by induction of cell death in cultured tobacco cells treated with tunicamycin. *Planta* **220**, 804-807.
- Iwata, Y., and Koizumi, N.** (2005b). An *Arabidopsis* transcription factor, AtbZIP60, regulates the endoplasmic reticulum stress response in a manner unique to plants. *Proc Natl Acad Sci U S A* **102**, 5280-5285.

- Jaquinod, M., Villiers, F., Kieffer-Jaquinod, S., Hugouvieux, V., Bruley, C., Garin, J., and Bourguignon, J.** (2007). A proteomics dissection of *Arabidopsis thaliana* vacuoles isolated from cell culture. *Mol Cell Proteomics* **6**, 394-412.
- Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W.** (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *Embo J* **6**, 3901-3907.
- Jelitto-Van Dooren, E.P., Vidal, S., and Denecke, J.** (1999). Anticipating endoplasmic reticulum stress. A novel early response before pathogenesis-related gene induction. *Plant Cell* **11**, 1935-1944.
- Johansen, L.K., and Carrington, J.C.** (2001). Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiol* **126**, 930-938.
- Johnson-Brousseau, S.A., and McCormick, S.** (2004). A compendium of methods useful for characterizing *Arabidopsis* pollen mutants and gametophytically-expressed genes. *Plant J* **39**, 761-775.
- Johnston, C.A., Temple, B.R., Chen, J.G., Gao, Y., Moriyama, E.N., Jones, A.M., Siderovski, D.P., and Willard, F.S.** (2007). Comment on "A G protein coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid". *Science* **318**, 914; author reply 914.
- Jones, P.R., Manabe, T., Awazuhara, M., and Saito, K.** (2003). A new member of plant CS-lyases. A cystine lyase from *Arabidopsis thaliana*. *J Biol Chem* **278**, 10291-10296.
- Josefsson, L.G., and Rask, L.** (1997). Cloning of a putative G-protein-coupled receptor from *Arabidopsis thaliana*. *Eur J Biochem* **249**, 415-420.
- Kader, J.C.** (1996). Lipid-Transfer Proteins in Plants. *Annu Rev Plant Physiol Plant Mol Biol* **47**, 627-654.
- Kalatzis, V., Cherqui, S., Antignac, C., and Gasnier, B.** (2001). Cystinosin, the protein defective in cystinosis, is a H(+)-driven lysosomal cystine transporter. *Embo J* **20**, 5940-5949.
- Kalatzis, V., Nevo, N., Cherqui, S., Gasnier, B., and Antignac, C.** (2004). Molecular pathogenesis of cystinosis: effect of CTNS mutations on the transport activity and subcellular localization of cystinosin. *Hum Mol Genet* **13**, 1361-1371.
- Kalatzis, V., Cohen-Solal, L., Cordier, B., Frishberg, Y., Kemper, M., Nuutinen, E.M., Legrand, E., Cochat, P., and Antignac, C.** (2002). Identification of 14 novel CTNS mutations and characterization of seven splice site mutations associated with cystinosis. *Hum Mutat* **20**, 439-446.
- Kamauchi, S., Nakatani, H., Nakano, C., and Urade, R.** (2005). Gene expression in response to endoplasmic reticulum stress in *Arabidopsis thaliana*. *Febs J* **272**, 3461-3476.
- Kanehisa, M., and Goto, S.** (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30.
- Keurentjes, J.J., Fu, J., Terpstra, I.R., Garcia, J.M., van den Ackerveken, G., Snoek, L.B., Peeters, A.J., Vreugdenhil, D., Koornneef, M., and Jansen, R.C.** (2007). Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci U S A* **104**, 1708-1713.

- Kinoshita, T., and Inoue, N.** (2000). Dissecting and manipulating the pathway for glycosylphosphatidylinositol-anchor biosynthesis. *Current Opinion in Chemical Biology* **4**, 632-638.
- Koiwa, H., Li, F., McCully, M.G., Mendoza, I., Koizumi, N., Manabe, Y., Nakagawa, Y., Zhu, J., Rus, A., Pardo, J.M., Bressan, R.A., and Hasegawa, P.M.** (2003). The STT3a subunit isoform of the Arabidopsis oligosaccharyltransferase controls adaptive responses to salt/osmotic stress. *Plant Cell* **15**, 2273-2284.
- Koizumi, N., Ujino, T., Sano, H., and Chrispeels, M.J.** (1999). Overexpression of a gene that encodes the first enzyme in the biosynthesis of asparagine-linked glycans makes plants resistant to tunicamycin and obviates the tunicamycin-induced unfolded protein response. *Plant Physiol* **121**, 353-361.
- Kornfeld, R., and Kornfeld, S.** (1985). Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* **54**, 631-664.
- Kranz, C., Denecke, J., Lehrman, M.A., Ray, S., Kienz, P., Kreissel, G., Sagi, D., Peter-Katalinic, J., Freeze, H.H., Schmid, T., Jackowski-Dohrmann, S., Harms, E., and Marquardt, T.** (2001). A mutation in the human MPDU1 gene causes congenital disorder of glycosylation type If (CDG-If). *J Clin Invest* **108**, 1613-1619.
- Krieg, J., Glasner, W., Vicentini, A., Doucey, M.A., Loffler, A., Hess, D., and Hofsteenge, J.** (1997). C-Mannosylation of human RNase 2 is an intracellular process performed by a variety of cultured cells. *J Biol Chem* **272**, 26687-26692.
- Kuchitsu, K., Yazaki, Y., Sakano, K., and Shibuya, N.** (1997). Transient cytoplasmic pH change and ion fluxes through the plasma membrane in suspension-cultured rice cells triggered by N-acetylchitooligosaccharide elicitor. *Plant Cell Physiol* **38**, 1012-1018.
- Kultz, D.** (2005). Molecular and evolutionary basis of the cellular stress response. *Annu Rev Physiol* **67**, 225-257.
- Kyte, J., and Doolittle, R.F.** (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132.
- Lalanne, E., and Twell, D.** (2002). Genetic control of male germ unit organization in Arabidopsis. *Plant Physiol* **129**, 865-875.
- Lalanne, E., Honys, D., Johnson, A., Borner, G.H., Lilley, K.S., Dupree, P., Grossniklaus, U., and Twell, D.** (2004). SETH1 and SETH2, two components of the glycosylphosphatidylinositol anchor biosynthetic pathway, are required for pollen germination and tube growth in Arabidopsis. *Plant Cell* **16**, 229-240.
- Lao, S.H., Loutre, C., Brazier, M., Coleman, J.O., Cole, D.J., Edwards, R., and Theodoulou, F.L.** (2003). 3,4-Dichloroaniline is detoxified and exported via different pathways in Arabidopsis and soybean. *Phytochemistry* **63**, 653-661.
- Leborgne-Castel, N., Jelitto-Van Dooren, E.P., Crofts, A.J., and Denecke, J.** (1999). Overexpression of BiP in tobacco alleviates endoplasmic reticulum stress. *Plant Cell* **11**, 459-470.
- Lebrun-Garcia, A., Chiltz, A., Gout, E., Bligny, R., and Pugin, A.** (2002). Questioning the role of salicylic acid and cytosolic acidification in mitogen-activated protein kinase activation induced by cryptogin in tobacco cells. *Planta* **214**, 792-797.

- Lee, D.H., and Kang, S.G.** (2007). Characterization of phosphatidylinositol-glycan biosynthesis protein class F gene in rice. *DNA Seq*, 1.
- Lehle, L., Strahl, S., and Tanner, W.** (2006). Protein glycosylation, conserved from yeast to man: a model organism helps elucidate congenital human diseases. *Angew Chem Int Ed Engl* **45**, 6802-6818.
- Lehrman, M.A., and Zeng, Y.** (1989). Pleiotropic resistance to glycoprotein processing inhibitors in Chinese hamster ovary cells. The role of a novel mutation in the asparagine-linked glycosylation pathway. *J Biol Chem* **264**, 1584-1593.
- Lehrman, M.A., Zhu, X.Y., and Khounlo, S.** (1988). Amplification and molecular cloning of the hamster tunicamycin-sensitive N-acetylglucosamine-1-phosphate transferase gene. The hamster and yeast enzymes share a common peptide sequence. *J Biol Chem* **263**, 19796-19803.
- Lerouxel, O., Mouille, G., Andeme-Onzighi, C., Bruyant, M.P., Seveno, M., Loutelier-Bourhis, C., Driouich, A., Hofte, H., and Lerouge, P.** (2005). Mutants in DEFECTIVE GLYCOSYLATION, an Arabidopsis homolog of an oligosaccharyltransferase complex subunit, show protein underglycosylation and defects in cell differentiation and growth. *Plant J* **42**, 455-468.
- Liaud, M.F., Brandt, U., and Cerff, R.** (1995). The marine red alga *Chondrus crispus* has a highly divergent beta-tubulin gene with a characteristic 5' intron: functional and evolutionary implications. *Plant Mol Biol* **28**, 313-325.
- Liu, X., Yue, Y., Li, B., Nie, Y., Li, W., Wu, W.H., and Ma, L.** (2007). A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science* **315**, 1712-1716.
- Livingstone, C.D., and Barton, G.J.** (1993). Protein-Sequence Alignments - a Strategy for the Hierarchical Analysis of Residue Conservation. *Computer Applications in the Biosciences* **9**, 745-756.
- Loque, D., Ludewig, U., Yuan, L., and von Wiren, N.** (2005). Tonoplast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate NH₃ transport into the vacuole. *Plant Physiol* **137**, 671-680.
- Lukowitz, W., Nickle, T.C., Meinke, D.W., Last, R.L., Conklin, P.L., and Somerville, C.R.** (2001). Arabidopsis *cyt1* mutants are deficient in a mannose-1-phosphate guanylyltransferase and point to a requirement of N-linked glycosylation for cellulose biosynthesis. *Proc Natl Acad Sci U S A* **98**, 2262-2267.
- Ma, H., Yanofsky, M.F., and Meyerowitz, E.M.** (1990). Molecular cloning and characterization of GPA1, a G protein alpha subunit gene from Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **87**, 3821-3825.
- Maresova, L., and Sychrova, H.** (2006). Arabidopsis thaliana CHX17 gene complements the *kha1* deletion phenotypes in *Saccharomyces cerevisiae*. *Yeast* **23**, 1167-1171.
- Marschner, H.** (1995). Mineral nutrition of higher plants. (London: Academic Press).
- Martin, R.C., Mok, M.C., Habben, J.E., and Mok, D.W.** (2001). A maize cytokinin gene encoding an O-glucosyltransferase specific to cis-zeatin. *Proc Natl Acad Sci U S A* **98**, 5922-5926.

- Martinez, I.M., and Chrispeels, M.J.** (2003). Genomic analysis of the unfolded protein response in *Arabidopsis* shows its connection to important cellular processes. *Plant Cell* **15**, 561-576.
- Mason, M.G., and Botella, J.R.** (2000). Completing the heterotrimer: isolation and characterization of an *Arabidopsis thaliana* G protein gamma-subunit cDNA. *Proc Natl Acad Sci U S A* **97**, 14784-14788.
- Mason, M.G., and Botella, J.R.** (2001). Isolation of a novel G-protein gamma-subunit from *Arabidopsis thaliana* and its interaction with Gbeta. *Biochim Biophys Acta* **1520**, 147-153.
- McGowan-Jordan, J., Stoddard, K., Podolsky, L., Orrbine, E., McLaine, P., Town, M., Goodyer, P., MacKenzie, A., and Heick, H.** (1999). Molecular analysis of cystinosis: probable Irish origin of the most common French Canadian mutation. *Eur J Hum Genet* **7**, 671-678.
- Mega, T.** (2004). Conversion of the carbohydrate structures of glycoproteins in roots of *Raphanus sativus* using several glycosidase inhibitors. *J Biochem* **136**, 525-531.
- Mellquist, J.L., Kasturi, L., Spitalnik, S.L., and Shakin-Eshleman, S.H.** (1998). The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochemistry* **37**, 6833-6837.
- Mikkelsen, M.D., Naur, P., and Halkier, B.A.** (2004). *Arabidopsis* mutants in the C-S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. *Plant J* **37**, 770-777.
- Mockler, T.C., Michael, T.P., Priest, H.D., Shen, R., Sullivan, C.M., Givan, S.A., McEntree, C., Kay, S.A., Chory, J.** (2007). The Diurnal Project: Diurnal and Circadian Expression Profiling, Model-based Pattern Matching, and Promoter Analysis. *Cold Spring Harbor Symposia on Quantitative Biology* **72**.
- Molina, A., and Garcia-Olmedo, F.** (1997). Enhanced tolerance to bacterial pathogens caused by the transgenic expression of barley lipid transfer protein LTP2. *Plant J* **12**, 669-675.
- Molina, A., Segura, A., and Garcia-Olmedo, F.** (1993). Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. *FEBS Lett* **316**, 119-122.
- Moriyama, E.N., Strope, P.K., Opiyo, S.O., Chen, Z., and Jones, A.M.** (2006). Mining the *Arabidopsis thaliana* genome for highly-divergent seven transmembrane receptors. *Genome Biol* **7**, R96.
- Moss, J.M., Reid, G.E., Mullin, K.A., Zawadzki, J.L., Simpson, R.J., and McConville, M.J.** (1999). Characterization of a novel GDP-mannose:Serine-protein mannose-1-phosphotransferase from *Leishmania mexicana*. *J Biol Chem* **274**, 6678-6688.
- Mou, Z., Wang, X., Fu, Z., Dai, Y., Han, C., Ouyang, J., Bao, F., Hu, Y., and Li, J.** (2002). Silencing of phosphoethanolamine N-methyltransferase results in temperature-sensitive male sterility and salt hypersensitivity in *Arabidopsis*. *Plant Cell* **14**, 2031-2043.
- Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., and Kimura, T.** (2007). Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J Biosci Bioeng* **104**, 34-41.

- Neill, S.J., Desikan, R., Clarke, A., Hurst, R.D., and Hancock, J.T.** (2002). Hydrogen peroxide and nitric oxide as signalling molecules in plants. *J Exp Bot* **53**, 1237-1247.
- Nemhauser, J.L., Hong, F., and Chory, J.** (2006). Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. *Cell* **126**, 467-475.
- Neves, S.R., Ram, P.T., and Iyengar, R.** (2002). G protein pathways. *Science* **296**, 1636-1639.
- Nickle, T.C., and Meinke, D.W.** (1998). A cytokinesis-defective mutant of *Arabidopsis* (*cyt1*) characterized by embryonic lethality, incomplete cell walls, and excessive callose accumulation. *Plant J* **15**, 321-332.
- Nikiforova, V., Freitag, J., Kempa, S., Adamik, M., Hesse, H., and Hoefgen, R.** (2003). Transcriptome analysis of sulfur depletion in *Arabidopsis thaliana*: interlacing of biosynthetic pathways provides response specificity. *Plant J* **33**, 633-650.
- Nishizawa, A., Yabuta, Y., Yoshida, E., Maruta, T., Yoshimura, K., and Shigeoka, S.** (2006). *Arabidopsis* heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant J* **48**, 535-547.
- Noctor, G., Strohm, M., Jouanin, L., Kunert, K.J., Foyer, C.H., and Rennenberg, H.** (1996). Synthesis of Glutathione in Leaves of Transgenic Poplar Overexpressing [γ]-Glutamylcysteine Synthetase. *Plant Physiol* **112**, 1071-1078.
- Noh, S.J., Kwon, C.S., Oh, D.H., Moon, J.S., and Chung, W.I.** (2003). Expression of an evolutionarily distinct novel BiP gene during the unfolded protein response in *Arabidopsis thaliana*. *Gene* **311**, 81-91.
- Obata, T., Kitamoto, H.K., Nakamura, A., Fukuda, A., and Tanaka, Y.** (2007). Rice shaker potassium channel *OsKAT1* confers tolerance to salinity stress on yeast and rice cells. *Plant Physiol* **144**, 1978-1985.
- O'Connor, T.R., Dyreson, C., and Wyrick, J.J.** (2005). Athena: a resource for rapid visualization and systematic analysis of *Arabidopsis* promoter sequences. *Bioinformatics* **21**, 4411-4413.
- Page, R.** (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357-358.
- Palmer, C.P., Aydar, E., and Djamgoz, M.B.** (2005). A microbial TRP-like polycystic-kidney-disease-related ion channel gene. *Biochem J* **387**, 211-219.
- Pandey, S., and Assmann, S.M.** (2004). The *Arabidopsis* putative G protein-coupled receptor GCR1 interacts with the G protein alpha subunit GPA1 and regulates abscisic acid signaling. *Plant Cell* **16**, 1616-1632.
- Pelham, H.R., Hardwick, K.G., and Lewis, M.J.** (1988). Sorting of soluble ER proteins in yeast. *Embo J* **7**, 1757-1762.
- Perfus-Barbeoch, L., Jones, A.M., and Assmann, S.M.** (2004). Plant heterotrimeric G protein function: insights from *Arabidopsis* and rice mutants. *Curr Opin Plant Biol* **7**, 719-731.
- Perlin, D.S., Brown, C.L., and Haber, J.E.** (1988). Membrane potential defect in hygromycin B-resistant *pma1* mutants of *Saccharomyces cerevisiae*. *J Biol Chem* **263**, 18118-18122.

- Pinto, J.T., Krasnikov, B.F., and Cooper, A.J.** (2006). Redox-sensitive proteins are potential targets of garlic-derived mercaptocysteine derivatives. *J Nutr* **136**, 835S-841S.
- Ponting, C.P., Mott, R., Bork, P., and Copley, R.R.** (2001). Novel protein domains and repeats in *Drosophila melanogaster*: insights into structure, function, and evolution. *Genome Res* **11**, 1996-2008.
- Reyes, F., Marchant, L., Norambuena, L., Nilo, R., Silva, H., and Orellana, A.** (2006). AtUTr1, a UDP-glucose/UDP-galactose transporter from *Arabidopsis thaliana*, is located in the endoplasmic reticulum and up-regulated by the unfolded protein response. *J Biol Chem* **281**, 9145-9151.
- Rhee, S.Y., Beavis, W., Berardini, T.Z., Chen, G., Dixon, D., Doyle, A., Garcia-Hernandez, M., Huala, E., Lander, G., Montoya, M., Miller, N., Mueller, L.A., Mundodi, S., Reiser, L., Tacklind, J., Weems, D.C., Wu, Y., Xu, I., Yoo, D., Yoon, J., and Zhang, P.** (2003). The *Arabidopsis* Information Resource (TAIR): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. *Nucleic Acids Res* **31**, 224-228.
- Rice, P., Longden, I., and Bleasby, A.** (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276-277.
- Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B.** (2003). An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol Biol* **53**, 247-259.
- Rost, B., Casadio, R., Fariselli, P., and Sander, C.** (1995). Transmembrane helices predicted at 95% accuracy. *Protein Sci* **4**, 521-533.
- Roudier, F., Fernandez, A.G., Fujita, M., Himmelsbach, R., Borner, G.H., Schindelman, G., Song, S., Baskin, T.I., Dupree, P., Wasteneys, G.O., and Benfey, P.N.** (2005). COBRA, an *Arabidopsis* extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. *Plant Cell* **17**, 1749-1763.
- Rush, J.S., and Waechter, C.J.** (1995). Transmembrane movement of a water-soluble analogue of mannosylphosphoryldolichol is mediated by an endoplasmic reticulum protein. *J Cell Biol* **130**, 529-536.
- Rush, J.S., and Waechter, C.J.** (1998). Topological studies on the enzymes catalyzing the biosynthesis of Glc-P-dolichol and the triglucosyl cap of Glc3Man9GlcNAc2-P-P-dolichol in microsomal vesicles from pig brain: use of the processing glucosidases I/II as latency markers. *Glycobiology* **8**, 1207-1213.
- Rylott, E.L., Rogers, C.A., Gilday, A.D., Edgell, T., Larson, T.R., and Graham, I.A.** (2003). *Arabidopsis* mutants in short- and medium-chain acyl-CoA oxidase activities accumulate acyl-CoAs and reveal that fatty acid beta-oxidation is essential for embryo development. *J Biol Chem* **278**, 21370-21377.
- Sadreyev, R., and Grishin, N.** (2003). COMPASS: a tool for comparison of multiple protein alignments with assessment of statistical significance. *J Mol Biol* **326**, 317-336.
- Saier, M.H., Jr., Eng, B.H., Fard, S., Garg, J., Haggerty, D.A., Hutchinson, W.J., Jack, D.L., Lai, E.C., Liu, H.J., Nusinew, D.P., Omar, A.M., Pao, S.S.,**

- Paulsen, I.T., Quan, J.A., Sliwinski, M., Tseng, T.T., Wachi, S., and Young, G.B.** (1999). Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochim Biophys Acta* **1422**, 1-56.
- Saito, K.** (2004). Sulfur assimilatory metabolism. The long and smelling road. *Plant Physiol* **136**, 2443-2450.
- Salcedo, G., Sanchez-Monge, R., Barber, D., and Diaz-Perales, A.** (2007). Plant non-specific lipid transfer proteins: an interface between plant defence and human allergy. *Biochim Biophys Acta* **1771**, 781-791.
- Sambrook, J., and Russell, D.W.** (2001). *Molecular cloning : a laboratory manual*. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).
- Samson, F., Brunaud, V., Balzergue, S., Dubreucq, B., Lepiniec, L., Pelletier, G., Caboche, M., and Lecharny, A.** (2002). FLAGdb/FST: a database of mapped flanking insertion sites (FSTs) of *Arabidopsis thaliana* T-DNA transformants. *Nucleic Acids Res* **30**, 94-97.
- Schenk, B., Fernandez, F., and Waechter, C.J.** (2001a). The ins(ide) and out(side) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. *Glycobiology* **11**, 61R-70R.
- Schenk, B., Imbach, T., Frank, C.G., Grubenmann, C.E., Raymond, G.V., Hurvitz, H., Korn-Lubetzki, I., Revel-Vik, S., Raas-Rotschild, A., Luder, A.S., Jaeken, J., Berger, E.G., Matthijs, G., Hennet, T., and Aebi, M.** (2001b). MPDU1 mutations underlie a novel human congenital disorder of glycosylation, designated type If. *J Clin Invest* **108**, 1687-1695.
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P.** (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**, 5857-5864.
- Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W.B., Flugge, U.I., and Kunze, R.** (2003). ARAMEMNON, a novel database for *Arabidopsis* integral membrane proteins. *Plant Physiol* **131**, 16-26.
- Sembdner, G., Atzorn, R., and Schneider, G.** (1994). Plant hormone conjugation. *Plant Mol Biol* **26**, 1459-1481.
- Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M., and Goff, S.A.** (2002). A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* **14**, 2985-2994.
- Shakin-Eshleman, S.H., Spitalnik, S.L., and Kasturi, L.** (1996). The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency. *J Biol Chem* **271**, 6363-6366.
- Shang, J., Gao, N., Kaufman, R.J., Ron, D., Harding, H.P., and Lehrman, M.A.** (2007). Translation attenuation by PERK balances ER glycoprotein synthesis with lipid-linked oligosaccharide flux. *J Cell Biol* **176**, 605-616.
- Shank, K.J., Su, P., Brglez, I., Boss, W.F., Dewey, R.E., and Boston, R.S.** (2001). Induction of lipid metabolic enzymes during the endoplasmic reticulum stress response in plants. *Plant Physiol* **126**, 267-277.

- Shotelersuk, V., Larson, D., Anikster, Y., McDowell, G., Lemons, R., Bernardini, I., Guo, J., Thoene, J., and Gahl, W.A.** (1998). CTNS mutations in an American-based population of cystinosis patients. *Am J Hum Genet* **63**, 1352-1362.
- Shpakov, A.** (2007). Serpentine Type Receptors and Heterotrimeric G-Proteins in Yeasts: Structural-Functional Organization and Molecular Mechanisms of Action. *Journal of Evolutionary Biochemistry and Physiology* **43**, 1-25.
- Simillion, C., Vandepoele, K., Van Montagu, M.C., Zabeau, M., and Van de Peer, Y.** (2002). The hidden duplication past of *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **99**, 13627-13632.
- Smyth, D.R., Bowman, J.L., and Meyerowitz, E.M.** (1990). Early flower development in *Arabidopsis*. *Plant Cell* **2**, 755-767.
- Sottosanto, J.B., Gelli, A., and Blumwald, E.** (2004). DNA array analyses of *Arabidopsis thaliana* lacking a vacuolar Na⁺/H⁺ antiporter: impact of AtNHX1 on gene expression. *Plant J* **40**, 752-771.
- Sparvoli, F., Faoro, F., Daminati, M.G., Ceriotti, A., and Bollini, R.** (2000). Misfolding and aggregation of vacuolar glycoproteins in plant cells. *Plant J* **24**, 825-836.
- Spiro, R.G.** (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R-56R.
- Sriraman, R., Bardor, M., Sack, M., Vaquero, C., Faye, L., Fischer, R., Finnern, R., and Lerouge, P.** (2004). Recombinant anti-hCG antibodies retained in the endoplasmic reticulum of transformed plants lack core-xylose and core- α (1,3)-fucose residues. *Plant Biotechnol J* **2**, 279-287.
- Sticher, L., and Metraux, J.P.** (2000). Inhibitors of N-linked glycosylation induce systemic acquired resistance in cucumber. *Physiol Mol Plant P* **56**, 245-252.
- Strahl-Bolsinger, S., Gentzsch, M., and Tanner, W.** (1999). Protein O-mannosylation. *Biochim Biophys Acta* **1426**, 297-307.
- Strasser, R., Altmann, F., Mach, L., Glossl, J., and Steinkellner, H.** (2004). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking β 1,2-linked xylose and core α 1,3-linked fucose. *FEBS Lett* **561**, 132-136.
- Strasser, R., Schoberer, J., Jin, C., Glossl, J., Mach, L., and Steinkellner, H.** (2006). Molecular cloning and characterization of *Arabidopsis thaliana* Golgi α -mannosidase II, a key enzyme in the formation of complex N-glycans in plants. *Plant J* **45**, 789-803.
- Strasser, R., Bondili, J.S., Vavra, U., Schoberer, J., Svoboda, B., Glossl, J., Leonard, R., Stadlmann, J., Altmann, F., Steinkellner, H., and Mach, L.** (2007). A unique β 1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell* **19**, 2278-2292.
- Stroobants, A.K., Hettema, E.H., van den Berg, M., and Tabak, H.F.** (1999). Enlargement of the endoplasmic reticulum membrane in *Saccharomyces cerevisiae* is not necessarily linked to the unfolded protein response via Ire1p. *FEBS Lett* **453**, 210-214.
- Sung, D.Y., Vierling, E., and Guy, C.L.** (2001). Comprehensive expression profile analysis of the *Arabidopsis* Hsp70 gene family. *Plant Physiol* **126**, 789-800.

- Tang, X., Frederick, R.D., Zhou, J., Halterman, D.A., Jia, Y., and Martin, G.B.** (1996). Initiation of Plant Disease Resistance by Physical Interaction of AvrPto and Pto Kinase. *Science* **274**, 2060-2063.
- Taylor, M.A., Ross, H.A., McRae, D., Stewart, D., Roberts, I., Duncan, G., Wright, F., Millam, S., and Davies, H.V.** (2000). A potato alpha-glucosidase gene encodes a glycoprotein-processing alpha-glucosidase II-like activity. Demonstration of enzyme activity and effects of down-regulation in transgenic plants. *Plant J* **24**, 305-316.
- Temple, B.R., and Jones, A.M.** (2007). The plant heterotrimeric G-protein complex. *Annu Rev Plant Biol* **58**, 249-266.
- Thoene, J., Lemons, R., Anikster, Y., Mullet, J., Paelicke, K., Lucero, C., Gahl, W., Schneider, J., Shu, S.G., and Campbell, H.T.** (1999). Mutations of CTNS causing intermediate cystinosis. *Mol Genet Metab* **67**, 283-293.
- Thornecroft, D., Sherson, S.M., and Smith, S.M.** (2001). Using gene knockouts to investigate plant metabolism. *J Exp Bot* **52**, 1593-1601.
- Town, M., Jean, G., Cherqui, S., Attard, M., Forestier, L., Whitmore, S.A., Callen, D.F., Gribouval, O., Broyer, M., Bates, G.P., van't Hoff, W., and Antignac, C.** (1998). A novel gene encoding an integral membrane protein is mutated in nephropathic cystinosis. *Nat Genet* **18**, 319-324.
- Toyomura, T., Murata, Y., Yamamoto, A., Oka, T., Sun-Wada, G.H., Wada, Y., and Futai, M.** (2003). From lysosomes to the plasma membrane: localization of vacuolar-type H⁺-ATPase with the $\alpha 3$ isoform during osteoclast differentiation. *J Biol Chem* **278**, 22023-22030.
- Tulsiani, D.R., Harris, T.M., and Touster, O.** (1982). Swainsonine inhibits the biosynthesis of complex glycoproteins by inhibition of Golgi mannosidase II. *J Biol Chem* **257**, 7936-7939.
- Turner, R.J., Taylor, D.E., and Weiner, J.H.** (1997). Expression of *Escherichia coli* TehA gives resistance to antiseptics and disinfectants similar to that conferred by multidrug resistance efflux pumps. *Antimicrob Agents Chemother* **41**, 440-444.
- Urade, R.** (2007). Cellular response to unfolded proteins in the endoplasmic reticulum of plants. *Febs J* **274**, 1152-1171.
- Van Loon, L.C., and Van Strien, E.A.** (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol Mol Plant P* **55**, 85-97.
- Veach, Y.K., Martin, R.C., Mok, D.W., Malbeck, J., Vankova, R., and Mok, M.C.** (2003). O-glucosylation of cis-zeatin in maize. Characterization of genes, enzymes, and endogenous cytokinins. *Plant Physiol* **131**, 1374-1380.
- Vierstra, R.D.** (1993). Protein-Degradation in Plants. *Annu Rev Plant Phys* **44**, 385-410.
- Vitale, A.** (2001). Uncovering secretory secrets: inhibition of endoplasmic reticulum (ER) glucosidases suggests a critical role for ER quality control in plant growth and development. *Plant Cell* **13**, 1260-1262.
- Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D.** (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J* **33**, 949-956.
- von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M.J.** (1993). Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase I and is

- unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol* **102**, 1109-1118.
- Wang, D., Weaver, N.D., Kesarwani, M., and Dong, X.** (2005). Induction of protein secretory pathway is required for systemic acquired resistance. *Science* **308**, 1036-1040.
- Wang, P., and Heitman, J.** (2005). The cyclophilins. *Genome Biol* **6**, 226.
- Wang, W., Vinocur, B., Shoseyov, O., and Altman, A.** (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci* **9**, 244-252.
- Wang, X., Shi, X., Li, Z., Zhu, Q., Kong, L., Tang, W., Ge, S., and Luo, J.** (2006). Statistical inference of chromosomal homology based on gene colinearity and applications to Arabidopsis and rice. *BMC Bioinformatics* **7**, 447.
- Wang, Z.Y., and Tobin, E.M.** (1998). Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* **93**, 1207-1217.
- Wang, Z.Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, M.S., and Tobin, E.M.** (1997). A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* **9**, 491-507.
- Ware, F.E., and Lehrman, M.A.** (1996). Expression cloning of a novel suppressor of the Lec15 and Lec35 glycosylation mutations of Chinese hamster ovary cells. *J Biol Chem* **271**, 13935-13938.
- Ware, F.E., and Lehrman, M.A.** (1998). Expression cloning of a novel suppressor of the Lec15 and Lec35 glycosylation mutations of Chinese hamster ovary cells. *J Biol Chem* **273**, 13366.
- Ware, F.E., Vassilakos, A., Peterson, P.A., Jackson, M.R., Lehrman, M.A., and Williams, D.B.** (1995). The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. *J Biol Chem* **270**, 4697-4704.
- Watanabe, R., Murakami, Y., Marmor, M.D., Inoue, N., Maeda, Y., Hino, J., Kangawa, K., Julius, M., and Kinoshita, T.** (2000). Initial enzyme for glycosylphosphatidylinositol biosynthesis requires PIG-P and is regulated by DPM2. *Embo J* **19**, 4402-4411.
- Weiss, C.A., Huang, H., and Ma, H.** (1993). Immunolocalization of the G protein alpha subunit encoded by the GPA1 gene in Arabidopsis. *Plant Cell* **5**, 1513-1528.
- Weiss, C.A., Garnaat, C.W., Mukai, K., Hu, Y., and Ma, H.** (1994). Isolation of cDNAs encoding guanine nucleotide-binding protein beta-subunit homologues from maize (ZGB1) and Arabidopsis (AGB1). *Proc Natl Acad Sci U S A* **91**, 9554-9558.
- Wilkinson, J.E., Twell, D., and Lindsey, K.** (1997). Activities of CaMV 35S and nos promoters in pollen: implications for field release of transgenic plants. *J Exp Bot* **48**, 265-275.
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J.** (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* **2**, e718.

- Wirtz, M., and Hell, R.** (2006). Functional analysis of the cysteine synthase protein complex from plants: structural, biochemical and regulatory properties. *J Plant Physiol* **163**, 273-286.
- Wopereis, S., Lefeber, D.J., Morava, E., and Wevers, R.A.** (2006). Mechanisms in protein O-glycan biosynthesis and clinical and molecular aspects of protein O-glycan biosynthesis defects: a review. *Clin Chem* **52**, 574-600.
- Wu, S.J., Ding, L., and Zhu, J.K.** (1996). SOS1, a Genetic Locus Essential for Salt Tolerance and Potassium Acquisition. *Plant Cell* **8**, 617-627.
- Xie, D.X., Feys, B.F., James, S., Nieto-Rostro, M., and Turner, J.G.** (1998). COI1: an Arabidopsis gene required for jasmonate-regulated defense and fertility. *Science* **280**, 1091-1094.
- Xu, L., Liu, F., Lechner, E., Genschik, P., Crosby, W.L., Ma, H., Peng, W., Huang, D., and Xie, D.** (2002). The SCF(COI1) ubiquitin-ligase complexes are required for jasmonate response in Arabidopsis. *Plant Cell* **14**, 1919-1935.
- Yang, J., Yu, M., Liu, B., Fan, B., Zhu, M., Xiong, T., and Li, K.** (2005). Cloning and initial analysis of porcine MPDU1 gene. *Asian Austral J Anim* **18**, 1237-1241.
- Yeats, T.H., and Rose, J.K.C.** (2008). The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Science* **17**, 191-198.
- Zeng, Y., and Elbein, A.D.** (1995). UDP-N-acetylglucosamine:dolichyl-phosphate N-acetylglucosamine-1-phosphate transferase is amplified in tunicamycin-resistant soybean cells. *Eur J Biochem* **233**, 458-466.
- Zeng, Y.C., and Lehrman, M.A.** (1990). A block at Man5GlcNAc2-pyrophosphoryldolichol in intact but not disrupted castanospermine and swainsonine-resistant Chinese hamster ovary cells. *J Biol Chem* **265**, 2296-2305.
- Zhai, Y., Heijne, W.H., Smith, D.W., and Saier, M.H., Jr.** (2001). Homologues of archaeal rhodopsins in plants, animals and fungi: structural and functional predictions for a putative fungal chaperone protein. *Biochim Biophys Acta* **1511**, 206-223.
- Zhang, K., and Kaufman, R.J.** (2004). Signaling the unfolded protein response from the endoplasmic reticulum. *J Biol Chem* **279**, 25935-25938.
- Zhu, X., Zeng, Y., and Lehrman, M.A.** (1992). Evidence that the hamster tunicamycin resistance gene encodes UDP-GlcNAc:dolichol phosphate N-acetylglucosamine-1-phosphate transferase. *J Biol Chem* **267**, 8895-8902.
- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Gruissem, W.** (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol* **136**, 2621-2632.